

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

(without alignments)
17054.732 Million cell updates/sec

Perfect score: 2980
Sequence: 1 gtgcaccacgcgtrccgcg.....aaaaaaaaagcgcgccgcta 2980

Gapor 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 4109280

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

```

1:  gb_ba:*
2:  gb_hcg:*
3:  gb_in:*
4:  gb_om:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pr:*
10: gb_ro:*
11: gb_srs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_prt:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_srs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_hum:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pdn:*
35: em_hcg_rdn:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcg_hum:*
40: em_hcg_mus:*
41: em_hto_other:*

```

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query Match	Length	DB	ID	Description
1	2868	96.2	3109	9	AF479827	AF479827 Homo sapi
2	2754	92.4	3007	9	AF479826	AF479826 Homo sapi
3	2708	90.9	2720	9	HSMB05307	AB134275 Homo sapi
4	2576	86.4	2576	9	AB058714	AB058714 Homo sapi
5	2259	75.8	2385	6	AX165526	AX165526 Sequence
6	1953	65.6	2128	9	BC016681	BC016681 Homo sapi
7	969	32.5	1014	9	HSMB03233	AB183145 Homo sapi
8	961	32.2	2025	6	AX327995	AX327995 Sequence
9	961	32.2	2219	6	AX327993	AX327993 Sequence
10	818	27.4	1956	9	HSAB6701	AJ006701 Homo sapi
11	678.8	22.8	134792	9	AC020922	AC020922 Homo sapi
12	491.8	16.5	3156	9	AK074411	AK074411 Homo sapi
13	446	15.0	3933	3	AB014885	AB014885 Halocynthia
14	391.2	13.1	3059	3	AF316542	AF316542 Caenorhab
15	373.6	12.4	5609	3	AY060288	AY060288 Drosophill
16	371	12.5	38000	9	AC008974	AC008974 Homo sapi
17	366.4	12.3	226060	2	AC079583	AC079583 Mus muscu
18	314	10.3	210105	2	AC125648	AC125648 Rattus no
19	307.2	10.3	1873	9	BC024291	BC024291 Homo sapi
20	296.8	10.0	2333	9	AF020089	AF020089 Homo sapi
21	269	9.0	3529	9	AB0808047	AB0808047 Homo sapi
22	269	9.0	3609	9	AB049127	AB049127 Homo sapi
23	267.4	9.0	2462	6	AX399887	AX399887 Sequence
24	267.4	9.0	3326	9	AY057448	AY057448 Homo sapi
25	267.4	9.0	3312	6	AX305105	AX305105 Sequence
26	267.4	9.0	3392	6	AX305106	AX305106 Sequence
27	267.4	9.0	4917	9	AB058763	AB058763 Homo sapi
28	261.6	8.8	1594	6	AX056390	AX056390 Sequence
29	259.4	8.7	3170	6	AX305103	AX305103 Sequence
30	259.4	8.7	3250	6	AX305104	AX305104 Sequence
31	245	8.2	2650	10	AB020480	AB020480 Rattus no
32	245	8.2	4092	10	AF106937	AF106937 Rattus no
33	240.6	8.1	2112	6	AX375128	AX375128 Sequence
34	240.6	8.1	2112	6	AF387638	AF387638 Homo sapi
35	240.6	8.1	2222	6	AX375131	AX375131 Sequence
36	240.6	8.1	2701	9	BC0008771	BC0008771 Homo sapi
37	240.6	8.1	2946	9	HSSTPKEMK	UY97630 H.sapiens m
38	238.4	8.0	2572	10	RNMARR2	Z83869 R.norvegicu
39	228.8	8.0	1749	3	AF020310	AF020310 Drosophill
40	227.2	7.7	2763	10	MMKEM	X70764 M.musculus
41	227.8	7.6	2681	3	AF181649	AF181649 Drosophal
42	225.4	7.6	1869	8	ATSKRIN2	X94757 A.thailana
43	225.4	7.6	2209	8	AY093170	AY093170 Arabidops
44	224.2	7.5	4498	10	MM011494	U11494 Mus muscullu
45	221.4	7.4	2352	6	AX262512	AX262512 Sequence


```

1890 CCAGCCCCGGGGGCGTGGGGGAGCCGCGGAGAGTCTGCTCACTCCATCCGCA 1949
1758 AAGAGCTCTGGGCTCCCGCTTCAACCGGGGCAAGAGAGGAGGCTTACGCTGAGG 1817
1950 AAGAGCTCTGGGCTCCCGCTTCAACCGGGGCAAGAGAGGAGGCTTACGCTGAGG 2009
1818 AAGAGCTCTGGGCTCCCGCTTCAACCGGGGCAAGAGAGGAGGCTTACGCTGAGG 1877
2010 AAGAGCTCTGGGCTCCCGCTTCAACCGGGGCAAGAGAGGAGGCTTACGCTGAGG 2069
1878 ACTTCATCTCTGTCGACAAAGAAATATCTCTGTCGTCATAGAGACAACTCTCA 1937
2070 ACTTCATCTCTGTCGACAAAGAAATATCTCTGTCGTCATAGAGACAACTCTCA 2129
1938 GAGAGATCAAGAGACATCTGTCATGCTTCTGTCGTCATGCTGTCATGTCATG 1997
2130 GAGAGATCAAGAGACATCTGTCATGCTTCTGTCGTCATGCTGTCATGTCATG 2189
1998 TGTCTGACAGACACAGCTTCAAGGCGGAGTCAAGAGGAGTGGCGGCTTCTCTCC 2057
2190 TGTCTGACAGACACAGCTTCAAGGCGGAGTCAAGAGGAGTGGCGGCTTCTCTCC 2249
2058 AAAAGCGCTCGCTTCAAGGCGGAGTCAAGAGGAGTGGCGGCTTCTCTCC 2117
2250 AAAAGCGCTCGCTTCAAGGCGGAGTCAAGAGGAGTGGCGGCTTCTCTCC 2309
2118 GAGGAGGAGGAGGAGGAGGAGTGGAGTCAAGGAGGAGTGGAGTGGAGTGGAGT 2177
2310 GAGGAGGAGGAGGAGGAGGAGTGGAGTCAAGGAGGAGTGGAGTGGAGTGGAGT 2369
2178 CAGAGCGTGGTCAAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2237
2370 CAGAGCGTGGTCAAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2429
2238 ACCAGCGCTCGCTTCAAGGCGGAGTCAAGAGGAGGAGTGGCGGCTTCTCTCC 2297
2430 ACCAGCGCTCGCTTCAAGGCGGAGTCAAGAGGAGGAGTGGCGGCTTCTCTCC 2489
2298 GAGGAGGAGGAGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2357
2490 GAGGAGGAGGAGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2549
2358 CAGGAGGAGGAGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2417
2550 CAGGAGGAGGAGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2609
2418 CAGGAGGAGGAGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2477
2610 CAGGAGGAGGAGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2669
2478 AAGAGCTCTGGGCTCCCGCTTCAACCGGGGCAAGAGAGGAGGCTTACGCTGAGG 2537
2670 AAGAGCTCTGGGCTCCCGCTTCAACCGGGGCAAGAGAGGAGGCTTACGCTGAGG 2729
2538 CAGGAGGAGGAGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2597
2730 CAGGAGGAGGAGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2789
2598 GAGAGGAGGAGGAGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2657
2790 GAGAGGAGGAGGAGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2849
2658 AAGAGCTCTGGGCTCCCGCTTCAACCGGGGCAAGAGAGGAGGCTTACGCTGAGG 2717
2850 AAGAGCTCTGGGCTCCCGCTTCAACCGGGGCAAGAGAGGAGGCTTACGCTGAGG 2909
2718 AAGAGCTCTGGGCTCCCGCTTCAACCGGGGCAAGAGAGGAGGCTTACGCTGAGG 2777
2910 AAGAGCTCTGGGCTCCCGCTTCAACCGGGGCAAGAGAGGAGGCTTACGCTGAGG 2969
2778 CAGGAGGAGGAGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2837

```

```

Db 2970 CTCCTCCCTGCTCTCCCGGACGCTTCTGTCAGGATTTCTCTCCGGAAGATTC 3029
Qy 2838 TGGTTTCGGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2897
Db 3030 TGGTTTCGGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3089
Qy 2898 AAAAAAAAAAAAAAAAAA 2917
Db 3090 AAAAAAAAAAAAAAAAAA 3109

RESULT 2
AF479826 3007 bp mRNA linear PRI 12-MAR-2002
LOCUS AF479826 complete cds.
DEFINITION Homo sapiens putative serine/threonine protein kinase mRNA.
ACCESSION AF479826
VERSION AF479826.1 GI:19401870
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3007)
AUTHORS She,X.Y., Guo,J.H. and Yu,L.
TITLE Direct Submission
JOURNAL Submitted (02-FEB-2002) School of Life Sciences, Laboratory of
Human Genes Research, Institute of Genetics, Fudan University, 220
Handan Road, Shanghai 200433, P. R. China
FEATURES
Source location/Qualifiers
1..3007
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13"
/issue.type="brain"
128..2512
/codon_start=1
/product="putative serine/threonine protein kinase"
/protein_id="FAL87697.1"
/db_xref="GI:19401871"
/translation="MVAGLLTGKSPSPDGVSPERKDEAVAGGSGEPEAEERGRHA
OYVPRLEKTKGKGTGKGLVKGHCITGCKVAKIKLVNREKLSVLMKEKRETAIIK
LIEHPVILKLDHYENKRIYILVLEHVSSEGLDYLYKKRLIPKARKPQVLSAL
DECHSYISICHRDLPENLIDERNRIADFGASLDVGSLSLETSCSPHYACPEYI
KBEKXGRRADWMSGVILFALVGLPEFDDNLROLLEVKRQVFMHPRTIPDCOS
LIRGMLIEVPEKRLSEIOIKHPWYLGKHEPDLCEPARGRYAMSLPSNELDPO
VLESMASLIGCFRDLRLHRLRSEENOEKMIYLLDREKRYSCADODLPPRNDVD
PPKRVDSMWLSHRGRPRPKRMEVSLTDAGGGSPPVTRRLLEMAHOSRBSVS
GASTGSSSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
PPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
SRNLSIRNLSIRNLSIRNLSIRNLSIRNLSIRNLSIRNLSIRNLSIRNLSIRNLSIRNLS
FLYLKDKPLSLISLADIVHAFSLISLHVSLSQTSFAEYKASGSPVFOKPVRFQYD
ISSSEBPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
LADENKAGTRPAGPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
BASE COUNT 636 a 940 c 906 g 525 t
ORIGIN
Query Match 92.4%; Score 2754; DB 9; Length 3007;
Best local Similarity 100.0%; Pred. No. 0;
Matches 2754; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 164 CAGCCCAATATGTGGCCCTTATCGGCTGGAAGAGAGCTGGGCAAGAGAGAGG 223
Db 254 CAGCCCAATATGTGGCCCTTATCGGCTGGAAGAGAGCTGGGCAAGAGAGAGG 313
Qy 224 CTGTTAACTCGGGGTCATCATCAGGCTCAGAGAGTGCATCAGATCAGTGAAC 283
Db 314 CTGTTAACTCGGGGTCATCATCAGGCTCAGAGAGTGCATCAGATCAGTGAAC 373
Qy 284 CAGGAGAGGAGTGCAGAGTGCAGTGAAGTGAAGAGGAGAGATCCATCTGAG 343
Db 374 CAGGAGAGGAGTGCAGAGTGCAGTGAAGTGAAGAGGAGAGATCCATCTGAG 433

```

QY 344 CTCATGAACACCCACATGTCCTCAAGCTCCACAGCTCTACGAGAACAGAAATATTG 403
 DB 434 CTCATGAACACCCACATGTCCTCAAGCTCCACAGCTCTACGAGAACAGAAATATTG 493
 QY 404 TACCTGTTCTGGAGACGCTCGGGGGGTGAGCTATTGCACTCTGTTAAAGAGGG 463
 DB 494 TACCTGTTCTGGAGACGCTCGGGGGGTGAGCTATTGCACTCTGTTAAAGAGGG 553
 QY 464 AGACTGACCCCAAGAGAGGCCGAAGTTCTCCGAGATTGTGTGCTGAGCTTC 523
 DB 554 AGACTGACCCCAAGAGAGGCCGAAGTTCTCCGAGATTGTGTGCTGAGCTTC 613
 QY 524 TCCACAGTACTCTCCATCTGCAAGAGACCTAAAGCCGAGAACTCTTTGATGAG 583
 DB 614 TCCACAGTACTCTCCATCTGCAAGAGACCTAAAGCCGAGAACTCTTTGATGAG 673
 QY 584 AAAAACACATCGCATGTGAGACTTCGGCATGCGCATGCGCTGAGAGGGGAGACCTC 643
 DB 674 AAAAACACATCGCATGTGAGACTTCGGCATGCGCATGCGCTGAGAGGGGAGACCTC 733
 QY 644 CTGAGACACGCTGCGGTCCTCCCATTTATGCTGTCCAGAGGTATTAAAGGGGAAAA 703
 DB 734 CTGAGACACGCTGCGGTCCTCCCATTTATGCTGTCCAGAGGTATTAAAGGGGAAAA 793
 QY 704 TATGATGCGCGCGGGGACACATGTGAGCTGTGAGTCACTCTTCGCTGCTGCTG 763
 DB 794 TATGATGCGCGCGGGGACACATGTGAGCTGTGAGTCACTCTTCGCTGCTGCTG 853
 QY 764 GGGGCTGCTCCCTTTGATGACGACACACTCCGCGCATGCTGAGAGAAAGGAAAGGGG 823
 DB 854 GGGGCTGCTCCCTTTGATGACGACACACTCCGCGCATGCTGAGAGAAAGGAAAGGGG 913
 QY 824 GTCCTTCACATGCCCCACCTTCCTCAGATTGCGCAGAGCTCTGAGGGGAAATGATC 883
 DB 914 GTCCTTCACATGCCCCACCTTCCTCAGATTGCGCAGAGCTCTGAGGGGAAATGATC 973
 QY 884 GAAGTGGAGCGCGGAAAAAGGCTCAGTCTGAGCAAAATTAGAAACATCTTGGTACTTA 943
 DB 974 GAAGTGGAGCGCGGAAAAAGGCTCAGTCTGAGCAAAATTAGAAACATCTTGGTACTTA 1033
 QY 944 GGGGGAACACAGACGACAGACCGCTGAGAGCCGCTGAGAGCCGCTGGCGGGTAAAGCATG 1003
 DB 1034 GGGGGAACACAGACGACAGACCGCTGAGAGCCGCTGAGAGCCGCTGGCGGGTAAAGCATG 1093
 QY 1004 CGGAGCCCTGCTCCAGAGGAGAGCTGAGACCCGAGCTCTTGAAGAGCATGSCATCATG 1063
 DB 1094 CGGAGCCCTGCTCCAGAGGAGAGCTGAGACCCGAGCTCTTGAAGAGCATGSCATCATG 1153
 QY 1064 GGCCTGCTCAGGAGACCGGAGAGGCTGATGCGAGCTGCGAGTGAAGAGAGAGAACCA 1123
 DB 1154 GGCCTGCTCAGGAGACCGGAGAGGCTGATGCGAGCTGCGAGTGAAGAGAGAGAACCA 1213
 QY 1124 GAAAGATGATATATTATCTGCTTTTGGATCGGAAGAGCGGTATCCAGCTGTGAGAC 1183
 DB 1214 GAAAGATGATATATTATCTGCTTTTGGATCGGAAGAGCGGTATCCAGCTGTGAGAC 1273
 QY 1184 CAGAGCCGCTCCCGGAAATGATGTTGAGACCCCGGAGAGCTGTGATTTCCCATG 1243
 DB 1274 CAGAGCCGCTCCCGGAAATGATGTTGAGACCCCGGAGAGCTGTGATTTCCCATG 1333
 QY 1244 CTGAGCCGCTCAGGAGACCGGAGACGAGAGGGAAGTCCATGGAATCTGAGCATCAC 1303
 DB 1334 CTGAGCCGCTCAGGAGACCGGAGACGAGAGGGAAGTCCATGGAATCTGAGCATCAC 1393
 QY 1304 GATCCGCGGGGTGCTGCTCCCTGTATCCACCCGAGCGGCTTGGAGATGGCCAGCAC 1363
 DB 1394 GATCCGCGGGGTGCTGCTCCCTGTATCCACCCGAGCGGCTTGGAGATGGCCAGCAC 1453
 QY 1364 AGCCAGAGATCCCGTAGAGTGAAGTGAAGCTTCACAGGGGTCTGCTCAGCCCTTAAGC 1423
 DB 1454 AGCCAGAGATCCCGTAGAGTGAAGTGAAGCTTCACAGGGGTCTGCTCAGCCCTTAAGC 1513
 QY 1424 AGCCAGAGATCCCGTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1483
 DB 1514 AGCCAGAGATCCCGTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1573
 QY 1484 GGGGGGGTCCCGGACTTCGAGAGAGCTGCTCTGAGAGAGCTGCTGAGAGAGAGAGAG 1543
 DB 1574 GGGGGGGTCCCGGACTTCGAGAGAGCTGCTCTGAGAGAGCTGCTGAGAGAGAGAGAG 1633
 QY 1544 GGGGGGGGAG 1603
 DB 1634 GGGGGGGGAG 1693
 QY 1604 GGCCTCCCGGCTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1663
 DB 1694 GGCCTCCCGGCTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1753
 QY 1664 AGTCCACCGGAG 1723
 DB 1754 AGTCCACCGGAG 1813
 QY 1724 GCGGCTGAG 1783
 DB 1814 GCGGCTGAG 1873
 QY 1784 CACCGGCGCAAGATCAGAGTCCCTACCGCTGAGAGAGATGCTGAGAGAGAGAGAGAG 1843
 DB 1874 CACCGGCGCAAGATCAGAGTCCCTACCGCTGAGAGAGATGCTGAGAGAGAGAGAGAG 1933
 QY 1844 TCCCGGAGCTGAG 1903
 DB 1934 TCCCGGAGCTGAG 1993
 QY 1904 CAATATCTCTGCTCTTAAAG 1963
 DB 1994 CAATATCTCTGCTCTTAAAG 2053
 QY 1964 GCTTTCTGCTGAGTCCGAGCTGAGTCAAGTGTGCTGACAGAGAGAGAGAGAGAGAG 2023
 DB 2054 GCTTTCTGCTGAGTCCGAGCTGAGTCAAGTGTGCTGACAGAGAGAGAGAGAGAGAG 2113
 QY 2024 GAGTACAG 2083
 DB 2114 GAGTACAG 2173
 QY 2084 ATCAGCTCTCTGAG 2143
 DB 2174 ATCAGCTCTCTGAG 2233
 QY 2144 ATCTACTCTGCTGAGTCTGAGTCAAGTGTGCTGAGAGAGAGAGAGAGAGAGAGAG 2203
 DB 2234 ATCTACTCTGCTGAGTCTGAGTCAAGTGTGCTGAGAGAGAGAGAGAGAGAGAGAG 2293
 QY 2204 GAGACATCTCAGGAG 2263
 DB 2294 GAGACATCTCAGGAG 2353
 QY 2264 GACGAG 2333
 DB 2354 GACGAG 2413
 QY 2324 CCACCGGCGCGGAG 2383
 DB 2414 CCACCGGCGCGGAG 2473
 QY 2384 AAGAGCTCTGAG 2443
 DB 2474 AAGAGCTCTGAG 2533
 QY 2444 GGGAG 2503
 DB 2534 GGGAG 2593
 QY 2504 AGGAGATGTCGGAG 2563

AUTHORS Nagase, T., Nakayama, M., Nakajima, D., Kikuno, R. and Ohara, O.
TITLE The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro

JOURNAL DNA Res. 8 (2), 85-95 (2001)
MEDLINE 21245130
REFERENCE 2 (bases 1 to 2576)
AUTHORS Ohara, O., Nagase, T. and Kikuno, R.

TITLE Direct Submission
JOURNAL Submitted (27-Mar-2001) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research, 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail: cdna@info.kazusa.or.jp,
URL: http://www.kazusa.or.jp/huge, Tel: 81-438-52-3913,
Fax: 81-438-52-3914)

FEATURES
SOURCE Location/Qualifiers
1..2576

1..2576
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="ha06731"
/tissue="brain"
/note="Vector: pBluescriptII SK plus"
1..2576
/gene="KIAA1811"
1..2149
/gene="KIAA1811"
/note="Start codon is not identified."
/codon_start=2
/product="KIAA1811 protein"
/protein_id="BAB47440.1"
/db_xref="GI:14017839"

BASE COUNT 508 a 842 c 766 g 460 t
ORIGIN
Query Match 86.4%; Score 2576; DB 9; Length 2576;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query 274 GATCGTGACCGGAGAGGCTGTCGAGTGGTCTGATGAGGTGAGCGGAGATGCG 333
Db 1 GATCGTGACCGGAGAGGCTGTCGAGTGGTCTGATGAGGTGAGCGGAGATGCG 60
Query 334 CATTCTGAACCTATGACACCAACCATGTCCTCAAGCTCCACGAGCTACGAGAACAA 393
Db 61 CATTCTGAACCTATGACACCAACCATGTCCTCAAGCTCCACGAGCTACGAGAACAA 120
Query 394 GAATATTTTACCTGTTCTGAGACAGTCTGGGGGGTGAAGTATTCGACTACCTGCT 453
Db 121 GAATATTTTACCTGTTCTGAGACAGTCTGGGGGGTGAAGTATTCGACTACCTGCT 180
Query 454 AAGAGAGGAGACTGACGCCCAAGAGAGGCGGAAATTTCTCCGCGAGATGTCTGCG 513
Db 181 AAGAGAGGAGACTGACGCCCAAGAGAGGCGGAAATTTCTCCGCGAGATGTCTGCG 240
Query 514 GCTGAGCTTGGCAGAGCTCTCATCTGCACAGAGACCTAAGACCGGAGACCTGCT 573
Db 241 GCTGAGCTTGGCAGAGCTCTCATCTGCACAGAGACCTAAGACCGGAGACCTGCT 300
Query 574 TTTGGATGAGAAAAACAACATCCGATTGACAGACTTGGCATGGCGTCCCTGAGTGG 633
Db 301 TTTGGATGAGAAAAACAACATCCGATTGACAGACTTGGCATGGCGTCCCTGAGTGG 360

Query 634 GGACAGCTCTCTGAGAGACCAAGCTGCGGTCCGCCATTATGCTGTCCAGAGGTGATTA 693
Db 361 GGACAGCTCTCTGAGAGACCAAGCTGCGGTCCGCCATTATGCTGTCCAGAGGTGATTA 420
Query 694 GGGGAAAAATATGATGCGCCCGCGGCAAGATGTGAGCTGTGAGATCTCTCTTCCG 753
Db 421 GGGGAAAAATATGATGCGCCCGCGGCAAGATGTGAGCTGTGAGATCTCTCTTCCG 480
Query 754 CTTGCTGCTGGGGGGCTCTGCTCTTGTATGACGACCAACCTCCGCGAGTGTGAGAGT 813
Db 481 CTTGCTGCTGGGGGGCTCTGCTCTTGTATGACGACCAACCTCCGCGAGTGTGAGAGT 540
Query 814 GAAAGGGGCGCTTTCACATGCCCCACCTTCATCTCCAGATTGCGAGAGCTCTGAG 873
Db 541 GAAAGGGGCGCTTTCACATGCCCCACCTTCATCTCCAGATTGCGAGAGCTCTGAG 600
Query 874 GGAATGATCGAAGTGGACCCCGAAAAAGGCTCAGCTGGACCAATTCAGAAACATCC 933
Db 601 GGAATGATCGAAGTGGACCCCGAAAAAGGCTCAGCTGGACCAATTCAGAAACATCC 660
Query 934 TTGCTACTAGGCGGAAACAGACGACCCGCTGCTGAGCCAGCCCTGGCCGCG 993
Db 661 TTGCTACTAGGCGGAAACAGACGACCCGCTGCTGAGCCAGCCCTGGCCGCG 720
Query 994 GGTAGCCTAGCGGAGCTCTGCTCCAAACGAGAGCTGACCCGAGCTCTAGAGAGCT 1053
Db 721 GGTAGCCTAGCGGAGCTCTGCTCCAAACGAGAGCTGACCCGAGCTCTAGAGAGCT 780
Query 1054 GGCATACAGGCGTCTGCTTCAGGGACCGGAGAGGCTCATCGGAGCTGGCGACTGAGA 1113
Db 781 GGCATACAGGCGTCTGCTTCAGGGACCGGAGAGGCTCATCGGAGCTGGCGACTGAGA 840
Query 1114 GGAGAACCAAGAAAGATGATATATATATCTGCTTTGGATCGAAGAGCGGTATCCAG 1173
Db 841 GGAGAACCAAGAAAGATGATATATATATCTGCTTTGGATCGAAGAGCGGTATCCAG 900
Query 1174 CTGTGAGACACAGGAGCTGCTCCCGGAAATGATGTTGACCCCCCGGAAAGCGTGGGA 1233
Db 901 CTGTGAGACACAGGAGCTGCTCCCGGAAATGATGTTGACCCCCCGGAAAGCGTGGGA 960
Query 1234 TTTTCCCATGCTGAGCGCTCAGCGGAGGAGGAGCGGACGAGCGGAGTCCATGGAAGTCT 1293
Db 961 TTTTCCCATGCTGAGCGCTCAGCGGAGGAGGAGCGGACGAGCGGAGTCCATGGAAGTCT 1020
Query 1294 GAGCATCACCGATGCGGCGGCTGCTCCCTGTACCCACCGAGCGGCTTGGAGAT 1353
Db 1021 GAGCATCACCGATGCGGCGGCTGCTCCCTGTACCCACCGAGCGGCTTGGAGAT 1080
Query 1354 GGCCCAAGCAGGAGGATCCGCTAGCGTCAAGGAGGCTCCAGGGTCTGTCTCCAG 1413
Db 1081 GGCCCAAGCAGGAGGATCCGCTAGCGTCAAGGAGGCTCCAGGGTCTGTCTCCAG 1140
Query 1414 CCTCTTAAGCAGCCCAAGAGATCCGCTCTTTTCCCTTTTACCGGAGCGGCGGCTGAGA 1473
Db 1141 CCTCTTAAGCAGCCCAAGAGATCCGCTCTTTTCCCTTTTACCGGAGCGGCGGCTGAGA 1200
Query 1474 TGAAGCTCGAGGCGGCGGCTCCCGGATCTCCAAAACGAGACGCTGCTTCTGGGGGCC 1533
Db 1201 TGAAGCTCGAGGCGGCGGCTCCCGGATCTCCAAAACGAGACGCTGCTTCTGGGGGCC 1260
Query 1534 CAGGGGTGGGGGCGCGGGGAGCAGCCCGCCCGAGTCCGCGTCCACACCCCTGCC 1593
Db 1261 CAGGGGTGGGGGCGCGGGGAGCAGCCCGCCCGAGTCCGCGTCCACACCCCTGCC 1320
Query 1594 CGGCGCCCAAGCTCCCGCGCTCTGAGCGGAGACCCCTTGGCATCTGCTGACAC 1653
Db 1321 CGGCGCCCAAGCTCCCGCGCTCTGAGCGGAGACCCCTTGGCATCTGCTGACAC 1380
Query 1654 GCCCGGGGCAAGTCCACACCGGAGACCCCGGGGACAAACACACCCCGGAGGCTGG 1713
Db 1381 GCCCGGGGCAAGTCCACACCGGAGACCCCGGGGACAAACACACCCCGGAGGCTGG 1440

QY	1714	CGTGGGGGAGACCCGCTGAGAGAGCGCTCAACCTCATCCGCAACACTTCCTGGGGCTC	1773
Db	1441	CGTGGGGGAGACCCGCTGAGAGACTCGCTCACTCCATCCGCAACACTTCCTGGGGCTC	1500
QY	1774	CCCTCCGCTTTCACCGGGCCAGATGACAGTCCCTACCGCTGAGGAGATGTCACAGTTGAC	1833
Db	1501	CCCTCCGCTTTCACCGGGCCAGATGACAGTCCCTACCGCTGAGGAGATGTCACAGTTGAC	1560
QY	1834	GCCAGACTCCGCCCCGGGAGCTGGGAAAAAGCTCCGTTGGGAACTTCATCTCTTGGG	1893
Db	1561	GCCAGACTCTCTCCCGGAGCTGGGAAAAAGCTCTCTGGTGGGAACTTCATCTCTTGGG	1620
QY	1894	CAAGAAGAACAAATATTTCTCGTGCCTAAAGACAAACTTCTGACGACATCAAGCAGA	1953
Db	1621	CAAGAAGAACAAATATTTCTCGTGCCTAAAGACAAACTTCTGACGACATCAAGCAGA	1680
QY	1954	CATGCTCCATGACCTTCTGTCGATGCCACAGCTGATGACAGTGTGCTGTACAGACCG	2013
Db	1681	CATGCTCCATGACCTTCTGTCGATGCCACAGCTGATGACAGTGTGCTGTACAGACCG	1740
QY	2014	CTTCAGGGCCGAGTACAAAGGCCAAGTGGCGCCCTCGCTCTCCAAAAGCCGACCGCTT	2073
Db	1741	CTTCAGGGCCGAGTACAAAGGCCAAGTGGCGCCCTCGCTCTCCAAAAGCCGACCGCTT	1800
QY	2074	CCAGGTGGAAATACGCTCTCTGTAGGGTCCAGAGCCCTCCCGGACGGGAGCGGACGCG	2133
Db	1801	CCAGGTGGAAATACGCTCTCTGTAGGGTCCAGAGCCCTCCCGGACGGGAGCGGACGCG	1860
QY	2134	AGGTGGTGGCATCTACTCCGTCACCTTCACCTCATCTCGGGTCCGAGCGCTGGTTCAA	2193
Db	1861	AGGTGGTGGCATCTACTCCGTCACCTTCACCTCATCTCGGGTCCGAGCGCTGGTTCAA	1920
QY	2194	GCGAGTGTGTGAGAACCATCCAGGCGACAGCTCTGTAGCAGCTCATGACCAAGCCCTCCGTGCA	2253
Db	1921	GCGAGTGTGTGAGAACCATCCAGGCGACAGCTCTGTAGCAGCTCATGACCAAGCCCTCCGTGCA	1980
QY	2254	GGGCTCTGGCAGCAGAGAAAGAGGGGGCCAGACCCGGCCCTGTCGTGGGCCCAACCCGAG	2313
Db	1981	GGGCTCTGGCAGCAGAGAAAGAGGGGGCCAGACCCGGCCCTGTCGTGGGCCCAACCCGAG	2040
QY	2314	CGTGCAGCCCCCAACCCGGCCCGCCAGACCCAGACTGAGCAGCTTCCCGCCGAGGCC	2373
Db	2041	CGTGCAGCCCCCAACCCGGCCCGCCAGACCCAGACTGAGCAGCTTCCCGCCGAGGCC	2100
QY	2374	CCCCAAGGCAAAACACTCTCTGGCCACCAAAGGGAGCCCTCTGCTGACCCCAAGGGGC	2433
Db	2101	CCCCAAGGCAAAACACTCTCTGGCCACCAAAGGGAGCCCTCTGCTGACCCCAAGGGGC	2160
QY	2434	CGGGAGGGAGGGGACCCCTTCACACCCCTTCCTGTGCCCCCAACTGTGCAATCTGTAA	2493
Db	2161	CGGGAGGGAGGGGACCCCTTCACACCCCTTCCTGTGCCCCCAACTGTGCAATCTGTAA	2220
QY	2494	ATTAAGGCCCAAGAAACATGTCCGGGAGGGGGGTGGACAAAAAACGGGCTTGCTTGACG	2553
Db	2221	ATTAAGGCCCAAGAAACATGTCCGGGAGGGGGGTGGACAAAAAACGGGCTTGCTTGACG	2280
QY	2554	GGATGGGGCTCCACAGGCGGTGCCCAACTGGGGGTGTTCTAGAGGAAACAGGGGGGGGG	2613
Db	2281	GGATGGGGCTCCACAGGCGGTGCCCAACTGGGGGTGTTCTAGAGGAAACAGGGGGGGGG	2340
QY	2614	GAGCGTTCCTATTTTATTTATTTGATTAATTTATTTATTTATTTATTTATTTGATCAATCTCT	2673
Db	2341	GAGCGTTCCTATTTTATTTATTTGATTAATTTATTTATTTATTTATTTATTTATTTGATCAATCTCT	2400
QY	2674	GGGGGTGGGGGGGGGGGAGGGAGCGGGAGCTGGTTGGGGGTGGCTTGAACATCCGACAG	2733
Db	2401	GGGGGTGGGGGGGGGGGAGGGAGCGGGAGCTGGTTGGGGGTGGCTTGAACATCCGACAG	2460
QY	2734	GCCCTCTGCTCCCTGTGTGTGTCCCAACCCCTCTTCCCGGGGCCCTGCTCCGCTGGTCT	2793
Db	2461	GCCCTCTGCTCCCTGTGTGTGTCCCAACCCCTCTTCCCGGGGCCCTGCTCCGCTGGTCT	2520
QY	2794	CCCCCAGACCTTGTGAAGGATTTGCTCTCCGAGAAAGAAATTTGCTTTCCCGTG	2849

Db	2521	CCCCCAGCAGCTTGTGTACGATTGCTCTCCGGAAGAAATCTGTTCGGTG	2576
RESULT 5			
LOCUS	AX166526	2385 bp	DNA
DEFINITION	Sequence 17 from Patent WO0138503.		linear
ACCESSION	AX166526		PAT 22-JUN-2001
VERSION	AX166526.1	GI:14546871	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 2385)		
TITLE	Plovanan,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,		
JOURNAL	Novel human protein kinases and protein kinase-like enzymes Patent: WO 0138503-A 17 31-May-2001;		
FEATURES	Sugen, Inc. (US)		
source	Location/Qualifiers		
BASE COUNT	1..2385		
ORIGIN	/organism="Homo sapiens" /db_xref="taxon:9606"		
Query Match	75.8%; Score 2259; DB 6; Length 2385;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches 2259; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
164	CACGCCCATATGTGGGCCCCCTATGTGGCTGTGGAAGAGCGTGGCCAAAGACAGACAGG	223	
127	CACGCCCAATATGTGGGCCCCCTATGTGGCTGTGGAAGAGCGTGGCCAAAGACAGACAGG	186	
224	CTGTGTTAACTGGGGTCCACTGCATGCATCAGCGGTGAGAAAGTGGCCATCAAGATCGTGAAC	283	
187	CTGTGTTAACTGGGGTCCACTGCATGCATCAGCGGTGAGAAAGTGGCCATCAAGATCGTGAAC	246	
284	CGGAGAGCTGTGGAGTGGTGTATGAAGGTGAGACGGAGATGCCATCTCTGAAG	343	
247	CGGAGAGCTGTGGAGTGGTGTATGAAGGTGAGACGGAGATGCCATCTCTGAAG	306	
344	CTCATCGAAGACCAACATGTCTCAAGTCCAGCGCTCAAGACGCTCAAGACAAAGAAATATTG	403	
307	CTCATCGAAGACCAACATGTCTCAAGTCCAGCGCTCAAGACGCTCAAGACAAAGAAATATTG	366	
404	TACGTGTTCTGGACAGCTCTGGGGGGTGAAGTATTCGACTACTGTGTAAGAAGGG	463	
367	TACGTGTTCTGGACAGCTCTGGGGGGTGAAGTATTCGACTACTGTGTAAGAAGGG	426	
464	AGACGTAGAGCCCAAGGAGGCCGAAAGTCTTCCGCCAGATTGTCTGCGCTGACTTC	523	
427	AGACGTAGAGCCCAAGGAGGCCGAAAGTCTTCCGCCAGATTGTCTGCGCTGACTTC	486	
524	TGCCACAGCTACTCATGTGACAGAGACCTAAAGCCCGAAGACCTGCTTGGATGAG	583	
487	TGCCACAGCTACTCATGTGACAGAGACCTAAAGCCCGAAGACCTGCTTGGATGAG	546	
584	AAAAACAACATCCGATTCGACAGACTTCGGCATGGCGTCCCTGCAAGGTGGGGACAAGCTC	643	
547	AAAAACAACATCCGATTCGACAGACTTCGGCATGGCGTCCCTGCAAGGTGGGGACAAGCTC	606	
644	CTGGAGACCTAGTGGGGTCCGCCCATTAATCGGTGTCAGAGGTGATTAAGGGGAAAAA	703	
607	CTGGAGACCTAGTGGGGTCCGCCCATTAATCGGTGTCAGAGGTGATTAAGGGGAAAAA	666	
704	TATGATGGCCCGGAGACATGTGAGAGCTGAGAGTCAATCCCTTGGCGCTGCTG	763	
667	TATGATGGCCCGGAGACATGTGAGAGCTGAGAGTCAATCCCTTGGCGCTGCTG	726	
764	GGGGCTCTGGCCCTTTGATGACGACACACTCCGACGCTCTGTGGAAGGTGAACGGGGC	823	

Db	1807	TCCCGGAGGCTGGCAAAACGCTCCGGTTCGGGAACTTCATCTCCTTGGACAAGAAGAA	1866
Qy	1904	CAATATTCCTCCTGGCTAAAGAGCAAACTCTCCAGACATCAAAAGCAGACATGTCAT	1963
Db	1867	CAATATTCCTGTCGTAAAGACAAACCTCTCCAGACATCAAAAGCAGACATGTCAT	1926
Qy	1964	GCTTTCCTGTGATCCCAAGCCTGAGTCACAGTGTGCTGCACAGACAGCTTCAGGGCC	2023
Db	1927	GCTTTCCTGTGATCCCAAGCCTGAGTCACAGTGTGCTGCACAGACAGCTTCAGGGCC	1986
Qy	2024	GAGTCAAGGCGCAGTGGGGGCCCCCTGCTTCCAAAAGCCGCTCCAGTTGAGTGCAC	2083
Db	1987	GAGTCAAGGCGCAGTGGGGGCCCCCTGCTTCCAAAAGCCGCTCCAGTTGAGTGCAC	2046
Qy	2084	ATCAGCTCCTCTGAGAGGTCCAGAGGCTCCCGCGCAGCAGGAGCGAGAGTGTGGC	2143
Db	2047	ATCAGCTCCTCTGAGAGGTCCAGAGGCTCCCGCGCAGCAGGAGCGAGAGTGTGGC	2106
Qy	2144	ATCTACTCCGTCACCTTACTCTCATCTTCGCGGTCCACAGCGTCGGTTCAAGCGAGTGGT	2203
Db	2107	ATCTACTCCGTCACCTTACTCTCATCTTCGCGGTCCACAGCGTCGGTTCAAGCGAGTGGT	2166
Qy	2204	GAGACATTCACAGGCAACACTCTCTGAGCACTCATGACACAGCCTCCGTCCAGGCCCTGGCA	2263
Db	2167	GAGACATTCACAGGCAACACTCTCTGAGCACTCATGACACAGCCTCCGTCCAGGCCCTGGCA	2226
Qy	2264	GAGGAGAAAGACGGGGGCCAGACCCGGCTGCTGGTGGCCCCACCCCGAAGCTGCAGCCC	2323
Db	2227	GAGGAGAAAGACGGGGGCCAGACCCGGCTGCTGGTGGCCCCACCCCGAAGCTGCAGCCC	2286
Qy	2324	CCACCCGGGGCCCGCAGAGCCCAAGCTGAGACACTCTCCCGCGCAGAGGCCCCCAAGGAC	2383
Db	2287	CCACCCGGGGCCCGCAGAGCCCAAGCTGAGACACTCTCCCGCGCAGAGGCCCCCAAGGAC	2346
Qy	2384	AAGAAGCTCTGGCCACCAAGGGAACCCCTCTGCCCTGA 2422	
Db	2347	AAGAAGCTCTGGCCACCAAGGGAACCCCTCTGCCCTGA 2385	

RESULT 6	BC016681	2128 bp	mRNA	linear	PRI 05-NOV-2001
LOCUS	BC016681				
DEFINITION	Homo sapiens, clone MGC:21870	IMAGE:4341838,	mRNA,	complete cds.	
ACCESSION	BC016681				
VERSION	BC016681.1	GI:16741775			
KEYWORDS	MGC.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 2128)				
AUTHORS	Strausberg,R.				
TITLE	Direct submission				
JOURNAL	Submitted (31-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov				
COMMENT	Contact: MGC help desk				

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lou Straut
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLN)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalobobcm.tmc.edu
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.

188

220

248

702

906

767

000

100

100

1

1
 2
 3

U
E
E

• • •

3

10

33

3

5

3

706

33

049

10

10

10

QY	1882	CATCTCTCTGGACAAAGAAGAAACAATATTCCTGCTCTAAAGACAAACTCTCAGAC	1941
Db	1	CATCTCTCTGGACAAAGAAGAAACAATATTCCTGCTCTAAAGACAAACTCTCAGAC	60
QY	1942	CATCAAGCAGACATGCGATCAGCTTCTCTGCGATTCGCCAGCCGAGCAGCAGTGTCT	2001
Db	61	CATCAAGCAGACATGCGATCAGCTTCTCTGCGATTCGCCAGCCGAGCAGCAGTGTCT	120
QY	2002	GTCCACGACCAAGCTTCAGGGCCGAGTACAAAGGCCAGTGGGGCCCTCCGTTCAAAA	2061
Db	121	GTCCACGACCAAGCTTCAGGGCCGAGTACAAAGGCCAGTGGGGCCCTCCGTTCAAAA	180
QY	2062	GCCCGTCCGCTTCCAGGTGTGACATACAGCTCTCTGAGGGTCCAGAGCCCTCCCGCGAG	2121
Db	181	GCCCGTCCGCTTCCAGGTGTGACATACAGCTCTCTGAGGGTCCAGAGCCCTCCCGCGAG	240
QY	2122	GGACGGCAGCGAGGGGGGATGATCATCTACCTCCTCAGCTTCATCTCAGTCCGATCAG	2181
Db	241	GGACGGCAGCGAGGGGGGATGATCATCTACCTCCTCAGCTTCATCTCAGTCCGATCAG	300
QY	2182	CCGTGCTTCAAGCGAGTGTGTGAGACCATCCAGGCACAGCTCTGAGCATCATGACCA	2241
Db	301	CCGTGCTTCAAGCGAGTGTGTGAGACCATCCAGGCACAGCTCTGAGCATCATGACCA	360
QY	2242	GCCCTTCCTGTGACAGCCCTGGCAGACAGAGAAGAAAGGGGGCCAGACCCGGCCGTGCTC	2301
Db	361	GCCCTTCCTGTGACAGCCCTGGCAGACAGAGAAGAAAGGGGGCCAGACCCGGCCGTGCTC	420
QY	2302	CCACCCCAAGACCTGTGACAGCCCAACCCGGCCGCGCAGACCCAGAGCTTAGCAGCTTCC	2361
Db	421	CCACCCCAAGACCTGTGACAGCCCAACCCGGCCGCGCAGACCCAGAGCTTAGCAGCTTCC	480
QY	2362	CCGCGCAGGCCCCCCCCAGGACAGAAAGCTCTGTGCCATCCAAAGGGACCCCTCTGCTCTG	2421
Db	481	CCGCGCAGGCCCCCCCCAGGACAGAAAGCTCTGTGCCATCCAAAGGGACCCCTCTGCTCTG	540
QY	2422	ACCCCAAGGGGGCGGGGAGGGAGGGAGCCCTCCACCCCTTCGTCGCCCCCAACT	2481
Db	541	ACCCCAAGGGGGCGGGGAGGGAGGGAGCCCTCCACCCCTTCGTCGCCCCCAACT	600
QY	2482	GTGAATCTGTAAATAAGAGCCCAAGGAACAATGTGCGGAGGGGGTGGACACAAAAACGGC	2541
Db	601	GTGAATCTGTAAATAAGAGCCCAAGGAACAATGTGCGGAGGGGGTGGACACAAAAACGGC	660
QY	2542	CTTGCCCTTCAGAGGATGGGGCTCCAGAGCCGTGCCCCAACTGGGGGTGGTCTAAGGGGA	2601
Db	661	CTTGCCCTTCAGAGGATGGGGCTCCAGAGCCGTGCCCCAACTGGGGGTGGTCTAAGGGGA	720
QY	2602	CAGGGGGCGGGGAGCGCTTCTAATTTATTTATTTATTTATTTATTTATTTATTTATTTG	2661
Db	721	CAGGGGGCGGGGAGCGCTTCTAATTTATTTATTTATTTATTTATTTATTTATTTATTTG	780
QY	2662	ATCAATCTCTGTCGGGGTGGGGTGGGGGAGGAGCGGAGCTGGTGGGGTGGCTTAGCA	2721
Db	781	ATCAATCTCTGTCGGGGTGGGGTGGGGGAGGAGCGGAGCTGGTGGGGTGGCTTAGCA	840
QY	2722	GATCCGAGACAGGGCCCTGTCCCTGTGTCGTCGCCAACCCCTCTTCCCGGGCCCTCC	2781
Db	841	GATCCGAGACAGGGCCCTGTCCCTGTGTCGTCGCCAACCCCTCTTCCCGGGCCCTCC	900
QY	2782	TGCCCTGTGTCCTCCCCCAGACACTTCTGTAGAGATTTGCTCTCCGGAAGGAATTCGAT	2841
Db	901	TGCCCTGTGTCCTCCCCCAGACACTTCTGTAGAGATTTGCTCTCCGGAAGGAATTCGAT	960
QY	2842	TTGCGCTGA 2850	

Db 961 TTCGCTGA 969

RESULT 8
AX327995
LOCUS AX327995 2025 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 3 from Patent WO0181358.
ACCESSION AX327995
VERSION AX327995.1 GI:18098148
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1
Meyers, R.
2246, protein kinase molecules and uses therefor
Patent: WO 0181588-A 3 01-NOV-2001;
JOURNAL Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1..2025
Source

BASE COUNT 442 a 641 c 616 g 326 t
ORIGIN

Query Match 32.2% Score 961; DB 6; Length 2025;
Best Local Similarity 70.1%; Pred. No. 4.7e-171;
Matches 1477; Conservative 0; Mismatches 435; Indels 195; Gaps 6;

159 CCCACGAGCCCAATATGTCGGCCCTATCGCTGGAGAAAGACGCTGGCAAGACAGA 218
29 CCGACGACGCGGAGATGTCGGCCCTACCGCTGGAGAAAGACGCTGGCAAGGCGAGA 88
219 CAGGCGCTGTTAACTCGGGGTCACATCGATCAGGCGTGAAGAGTCCCATCAAGATCG 278
89 CAGGCTGCTGTAAGTGGGGGTTCACTGCTCAGCTCCAGAAAGTGGCCATCAAGATCG 148
279 TGAACCGGAGAGCTGTCGAGTGGTCTGATGAAGGTGAGCGGAGATCGCCATCC 338
149 TCAACCTGAGAGGCTCAGCGAGTGGTCTGATGAAGGTGAGCGGAGATCGCCATCC 208
339 TGAAGCTATGAAACACCATGTCCTCAAGCTCCACAGACGCTCAAGAAACAAAT 398
209 TGAAGCTATGAGACACCCCGCTCAAGCTGACGACGCTTATGAACAAACAAAT 268
399 ATTTGTACCTGTTCTGAGACGCTCTCGGGGGTGAAGCTATTTGACCTGTTAAAG 458
269 ATTTGTACCTGTTCTGAGACGCTCTCGGGGGTGAAGCTATTTGACCTGTTAAAG 328
459 AGGGAGACGACGCGCCCAAGAGGCGCCCAAGATTTCTCCGCAAGATTTGTCGCTGG 518
329 AGGGAGGCTTACGCTTAAGGAGGCTCGAAGTTCTCCGGGACATCTCCGCTGG 388
519 ACTTCTGACAGCTACTCATCTGCAACAGAGACTTAAGGCCGGAACCTGCTTTGG 578
389 ACTTCTGACAGCTACTCATCTGCAACAGAGACTTAAGGCCGGAACCTGCTTTGG 448
579 ATGAGAAAAACAATCCGATTGCAAGCTTCGACGCTCCCTGACAGTGGGGGAGA 638
449 ACGAGAAAGAACATCCGATCGCAGACTTTGGCATGGCGTCCCTGCAAGTGGGGAGA 508
639 GCCCTCCGAGAACGACGCTGGGGTCCCGCATTTAGCGTTCACAGAGTGAAGGGG 698
509 GCCCTGTTGAGAACGACGCTGGGGTCCCGCATTTAGCGTTCACAGAGTGAAGGGG 568
699 AAAAATATGATGCGCGCGCGAGACATGAGCTGTGAGATCAATCCCTCGCCCTCG 758
569 AGAAGTATGACGGCGCGAGCGGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 628
759 TCGTGGGGGCTTCGCTTTGAGACGAAACCTCCGCGAGCTGTGAGAGTGAAG 818
629 TGGTGGGGGCTTCGCTTTGAGACGAAACCTCCGCGAGCTGTGAGAGTGAAG 688

819 GGGGGCTTCCACATGCCCCACTTCTCTCCAGATTGCGAGAGCTCTTGAGGGAA 878
689 GGGGGCTTCCACATGCCCCACTTCTCTCCAGATTGCGAGAGCTCTTGAGGGAA 748
879 TGATGAAATGAGAGCGCGAAAAAGGCTGAGTGGAGCAAAATTCAGAAACATCTTGGT 938
749 TGAGGAGGTGAGAGCGCGAGCGCGGCTCAGGCTAAGAGCAATTCAGAAACATATGTT 808
939 ACCTAGCGGGGAAACAGAGCGAGAGCGCTGCTGAGAGCGAGCGCGCGGGTAG 998
809 ATATAGGGGGCAAGATGAGCGCGGAAAC-----AGAGAGCCCATTTCTCGCAAGGTGC 862
999 CCATGCGGAGCGCTGCCATCCAGAGAGAGCTGAGAGCGCGCGCGCGCGAGAGATGAGAT 1058
863 AGATCCGCTCGCTGCCAGCGCTGAGAGAGATGAGAGCGCGCGCGCGCGAGATGAGAT 922
1059 CACTGGGCTGCTTCAAGGAGCGCGAGAGCGCTGAGAGAGCTGCGAGAGAGAGAGA 1118
923 CACTGGGCTGCTTCAAGGAGCGCGAGAGCGCTGAGAGAGCTGCGAGAGAGAGAGA 982
1119 ACCAAGAAAGATGATATATATATCTGCTTTGAGTGAAGAGCGGATCCAGCTGTG 1178
983 ACCAGAGAAAGATGATATATATCTGCTTTGAGTGAAGAGCGGATCCAGAGAGAG 1042
1179 AGGACGAGAGAGCTGCTCCCGGAAATGATGAGAGCGCGGAGAGCGGATGATTC 1238
1043 AGGATGAGAGAGCTGCTCCCGGAAATGATGAGAGCGGAGAGCGGATGATTC 1102
1239 CCATGCTGAGCGCTGAGAGAGCGCGGAGAGCGGAGAGCGGATGATTC 1298
1103 CGATGCTGAGCGCGCTGAGAGAGCGCGGAGAGCGGAGAGCGGATGATTC 1162
1299 TCACCGATGCGCGGAGAGCGCGGAGAGCGGAGAGCGGATGATTC 1358
1163 TGAC-----GGAGCGGCTGCTCCCGGAGAGCGGAGAGCGGATGATTC 1213
1359 AGCAGAGAGAGAGCTGCTCCCGGAAATGATGAGAGCGGAGAGCGGATGATTC 1418
1214 AGCAGAGAGAGAGCTGCTCCCGGAAATGATGAGAGCGGAGAGCGGATGATTC 1273
1419 TAAAGCGCCAGAGAGAGCGGAGAGCGGAGAGCGGAGAGCGGATGATTC 1478
1274 TCACCGAGCGCGGAGAGAGCGGAGAGCGGAGAGCGGATGATTC 1286
1479 CTGAGAGAGAGAGAGCGGAGAGCGGAGAGCGGAGAGCGGATGATTC 1538
1287 ----- 1286
1539 GTGGGGGCGCGGGGAGAGAGCGGAGAGCGGAGAGCGGAGAGCGGATGATTC 1598
1287 -----GGTGAACCTGACCGCTTCAACCAAGAGAGAGAGCGGATGATTC 1326
1599 CCGCAGAGCTGCGCGCGCTCTCTGAGAGAGCGGAGAGCGGATGATTC 1658
1327 -----ACCCCAAGAGAGAGAGAGCGGAGAGCGGATGATTC 1357
1659 GGGCAGTCCACCGGAGAGCGGAGAGCGGAGAGCGGAGAGCGGATGATTC 1718
1358 AGGAG 1411
1719 GGGAG 1778
1412 GAG 1471
1779 GCTTTCACCGGAG 1838
1472 GCTTTCACCGGAG 1531
1839 AGTCTCCCGGAG 1898
1532 AGTCTCCCGGAG 1591

1899 AAGAAATATTCCTGCTGCTAAGAGCAACCTCTGACGATCAAGACATCG 1958
 1592 AGGAGAGATCTTCTGCTGATCAACCAACCTCTGATCAAGCTGACATCA 1651
 1959 TCCATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2018
 1652 TCCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1711
 2019 GGGGAGATCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2078
 1712 GGGGAGATCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1771
 2079 TGGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2138
 1772 TGGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1813
 2139 GGGGAGATCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2198
 1814 AGGAGATCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1873
 2199 TGGGAGATCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2258
 1874 TGGGAGATCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1933
 2259 TGGGAGATCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2265
 1934 TGGGAGATCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1940

RESULT 9
 LOCUS AX327993 2219 bp DNA linear PAT 07-JAN-2002
 DEFINITION Sequence 1 from Patent WO0181588.
 ACCESSION AX327993
 VERSION AX327993.1 GI:18098146
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 AUTHORS Meyers, R.
 TITLE 2246, Protein kinase molecules and uses therefor
 JOURNAL Patent: WO 0181588-A 1 01-NOV-2001;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source location/Qualifiers
 1. 2219
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 53. 2077
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD20436.1"
 /db_xref="GI:18098147"
 /translation="MTSTGKDGAGNAQYVPRLETKGQYGLYKLVHYCTQCK
 VAIKIVNREKLSVLMKVEREIAIKLIEHPYKLDHYENKRYLYLVIEHVGGE
 LFDYLVKGRITPKEARKFERQJISALDFCHSHSICHARDIKPENMLIDENKNIAD
 GMAISQVDSLETSFGSPHYACPEVIRGKDYKRSKPLSPRGSPPLPKTPVH
 DNIRLEKVRKGVFHPHPIPDOSLIGMSVPSVPRVPRGSPPLPKTPVH
 PEPEPIPRKVRKSLPLEDIDPDVLSHSGSCPPDRKLLDILSEENOKMY
 FLIDRKERTPSOEDDLPPRNEIDPPRKVDSPMLNRHGRKREKSMELVSTYDGG
 SPVAPRAIEMAHQGRSISGASGLSTSPSPVPRVPRGSPPLPKTPVH
 TPESPAGTPNPPTSPSPSPVPGVPMARLMSIKNSPLGSPRHRKLOVPBEKSNL
 TPESPELAKSWFNIEISLEKEQIVVIKDRPLSIIKADIYAFILSISHSVIS
 OTSFREYKATGGAPEOKPFOVDITYTEGGAGKENGISYVFTILSGSPRFRK
 VETIOAOLSTHDPAPAOHLSPRPAPLSMAGLKGKQVATSYESSL"

CDS
 476 a 716 c 683 g 344 t

BASE COUNT
 ORIGIN

Query Match 32.2%; Score 961; DB 6; Length 2219;
 Best local Similarity 70.1%; Pred. No. 4,7e-171;
 Matches 1477; Conservative 0; Mismatches 435; Indels 195; Gaps 6;

159 CCCAGACGCCCAATATGTTGGCCCTTATCGCTGGAGAAAGACGCTGGCAAGGACAGA 218
 81 CCCAGACGCCCAATATGTTGGCCCTTATCGCTGGAGAAAGACGCTGGCAAGGACAGA 140
 219 CAGGCTGTGTTAAATCGGGGCTCAGTGCATGACGGGCTGAAAGGCTGCAATCAAGATCG 278
 141 CAGGCTGTGTTAAATCGGGGCTCAGTGCATGACGGGCTGAAAGGCTGCAATCAAGATCG 200
 279 TGAACCGGAGAGAGCTGTGAGTCTGCTGATGAGGTGAGAGCGGAGATCGCCATCC 338
 201 TCAACCGGAGAGAGCTGTGAGTCTGCTGATGAGGTGAGAGCGGAGATCGCCATCC 260
 339 TGAAGCTCATTCGACACCCACATGCTCTCAAGCTCCACGAGCTTACGAGAAACAAGAT 398
 261 TGAAGCTCATTCGACACCCACATGCTCTCAAGCTCCACGAGCTTACGAGAAACAAGAT 320
 399 ATTGTGACCTGCTGTGAGACAGCTGTGAGGCTGAGGCTGAGATTCGATCTGCTGTAAGA 458
 321 ATTGTGACCTGCTGTGAGACAGCTGTGAGGCTGAGGCTGAGATTCGATCTGCTGTAAGA 380
 459 AGGAGAGCTGACGCGCCCAAGAGAGGCGCCGAAAGTCTTCCGACAGATTTGTCTGCGCTGG 518
 381 AGGAGAGCTGACGCGCCCAAGAGAGGCGCTGAGAGGCTGAGAGTCTTCCGACAGATTCATCT 440
 519 ACTTGTCCACAGACTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 578
 441 ACTTGTCCACAGACTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 500
 579 ATGAG 638
 501 ATGAG 560
 639 GCTCTCTGAGAGACAGCTGCGGGTCCGCCATTAATGCTGCTGAGAGAGTGAAGGGGG 698
 561 GCTCTCTGAGAGACAGCTGCGGGTCCGCCATTAATGCTGCTGAGAGAGTGAAGGGGG 620
 699 AAAAATATGATGCG 758
 621 AGAAGATGATGCG 680
 759 TCGTGGGGGCTGCGCTTGTGATGACAGCAACCTCCGCAAGCTCTGAGAGAGTGAAG 818
 681 TCGTGGGGGCTGCGCTTGTGATGACAGCAACCTCCGCAAGCTCTGAGAGAGTGAAG 740
 819 GGGGCTCTTCCACATATGCGCCCACTTATTCCTGAGATTCGAGAGAGCTCTGAGAGG 878
 741 GGGGCTCTTCCACATATGCGCCCACTTATTCCTGAGATTCGAGAGAGCTCTGAGAGG 800
 879 TGAATCAAGTGAAGCG 938
 801 TGAATCAAGTGAAGCG 860
 939 ACTTGGCGGAG 998
 861 ATTTAGGGGCGAGAGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 914
 999 CCATGCGGAG 1058
 915 AGATCGCGCTGCG 974
 1059 CACTGGCTCTTCAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1118
 975 CACTGGCTCTTCAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1034
 1119 ACCAAGAGAGAGAGATATATATCTGCTTGTGATGAGAGAGAGAGAGAGAGAGAGAG 1178
 1035 ACCAAGAGAGAGATATATATCTGCTTGTGATGAGAGAGAGAGAGAGAGAGAGAGAG 1094
 1179 AGGAG 1238
 1095 AGGAG 1154
 1239 CCATGCTGAGCGCTGACGAG 1298

Db	1155	CGATGCTGAAACCCGGCACGGCAAGGCGGGCCAGAAACCAATTCATGGAGGTGCTCAAGC	1214
OY	1299	TCACCGATGCGGGGGGGTGGGTGCTCCCTCTACCCACCAGCGGGGCTTGGAGATGGCC	1358
Db	1215	TGAC-----GACGGGGGGCTCCCGGTGGCTGGCGGGGCCATTGATGAGTGGCC	1265
OY	1359	AGCAGCCAGAGATCCGATAGCGCTAGTGGAGCCCTCAGGGGTCTGTCCACAGCCCTC	1418
Db	1266	AGCAGCGCCAGAGGTCTCGGTCCATCAGCGGTGGCTCTCCTCAGGCTTTCCACAGCCAC	1325
OY	1419	TAAAGAGCCCAAGAGAGTCGGTCTTTTCTTTACCGGAGCCGGGGGCTTGGATGAGG	1478
Db	1326	TCAGCAGCCCCC-----	1338
OY	1479	CTTCAGGCGGGGGCTCCCGACTTCCAAAAAGCAGACGTGCTTCTCGGGGCCCAAGG	1538
Db	1339	-----	1338
OY	1539	GTGGGGGGCCCGGGAGAGCAAGCCCCCGCCCAAGTGGCCGCTCCACACCCCTGCCCCG	1598
Db	1339	-----GGTACCCCTCACCCCTCACCAAGGGGAGTCCCTCC-----	1378
OY	1599	CCCCAGGCTCCCGCGCTCCTTGGCGGAGCCCCCTTGCACTGGCCCTGTGACAGGCC	1658
Db	1379	-----ACCCCAAGGGAGACCTGTGCACAGCCAA	1409
OY	1659	GGGCGAGTCCACCGGAGACCGGGGACAACACACCCCGCAGCCCGCGGGCGTGG	1718
Db	1410	AGGAGAGCCCGGTGGCACGCCCAACCCACGCGCCGCTCAGCCCC-----AGCGTGG	1463
OY	1719	GGGAGACCGCGTGGAGAGTGCCTCAACACCAATCCGCAAGCTTCGTGGGCTCCCTC	1778
Db	1464	GAGGGTGGCTTGGAGGGCGCGGCTCAACTCACAAGAAAGCTTTCTGAGCTACCCC	1533
OY	1779	GCTTTCACGGGCAAGATGCAAGTCCATCCGCTGAGAGATGTCCAGTTGACGCGAG	1838
Db	1524	GCTTCACGCGCGAAACAGTCAAGTTCCGACGCGGAGGAGATGTCCAACTGACACAG	1583
OY	1839	AGTCCCTCCCGAGCGTGGGCAAAAGCTCCTGTTGGGAACTTCATCTCTTGGACAAAG	1898
Db	1584	AGTCGTCGCCAGAGCTGGCGAAGAAAGTCTGGTTGGAACTTCATCAGCTGGAGAAAG	1643
OY	1899	AAGAACAAATATTCTCTGTCCTAAAGAGCAAAACCTCTCAGCAGCATCAAGAAGACATCG	1958
Db	1644	AGGAGCAGATCTTGGTGCATCAAAAGCAAAACCTCTGAGCTGCATCAAGGCTGACATCG	1703
OY	1959	TCCATGCCCTTGTGTGATCCCAAGCTGAGTACAGTGTGCTGACAGACCAAGCTTCA	2018
Db	1704	TCCAGCCCTTGTGATTTCCCAAGTCAAGCCACAGCGCATCTCCAAACAGCTTCC	1763
OY	2019	GGGCGAGTACAAGGCAAGTGGCGGCGCCCTCCGCTTCCAAAAGCCGTCGGTCCAGG	2078
Db	1764	GGGCGAGTACAAGGCAAGGCGGGGGGCCAGCGGTGTTCCAGAAGCCGGTCAAGTTCCAGG	1823
OY	2079	TGGACATCAGCTCTCTYAGGGTCCAGAGCCCTCCCCGCGAGCGGAGCGGAGGAGTG	2138
Db	1824	TTGAATATCACTAACGAGGGT-----GGGAGGGCGCAAGAGAGA	1865
OY	2139	GTGGCATCTACCGGTCACCTTCACTCTCATCTCGGGGTCCAGCGGCTGTTAAACGAG	2198
Db	1866	ACGGCATCTACTCCGTCACTTCACTCTCTTCAAGGCCCGCAGCGCTGCTTAAAGAGGG	1925
OY	2199	TGTTGAGAACCATCCAGGCAAGCTCTGAGCACTTATGACAGGACCGCTCGTTCAGAGCC	2258
Db	1926	TGTTGAGAACCATCCAGGCGCCAGCTGTGTGAGCACAAGCAACCGCGCTGGGGCCACACT	1985
OY	2259	TGGCAGA 2265	
Db	1986	TGTCAGA 1992	

HSAs6701	1956 bp	mRNA	linear	PRI 14-MAR-2001
LOCUS				
DEFINITION	HsA6701	Homo sapiens mRNA for putative serine/threonine protein kinase,		
	partial.			
ACCESSION	AJ006701			
VERSION	AJ006701.1	GI:3217027		
KEYWORDS	putative; serine/threonine protein kinase.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 1956) Stanchi,F., Bertocco,E., Toppo,S., Dioguardi,R., Simonati,B., Cannata,N., Zimbello,R., Lanfranchi,G. and Valle,G.			
TITLE	Characterization of 16 novel human genes showing high similarity to Yeast sequences			
JOURNAL	Yeast 18 (1), 63-80 (2001)			
MEDLINE	21064499			
PUBMED	11124703			
REFERENCE	2 (bases 1 to 1956)			
AUTHORS	Stanchi,F.			
TITLE	Direct Submision			
JOURNAL	Submitted (02-JUN-1998) Stanchi F., CRIBI Biotechnology Centre, Universita' di Padova, Via G. Colombo 3, 35121 Padova, 35121, ITALY			
FEATURES	Location/Qualifiers			
source	1..1956			

```

CDS
      /clone_id="Soares infant brain IN1B"
      /dev_stage="73 days post natal"
      <1..1814
      /codon_start=3
      /product="putative serine/threonine protein kinase"
      /protein_id="CAA07196.1"
      /db_xref="GI:3217028"
      /db_xref="SPTREMBL:O60843"
      /translation="LIEHPVIVKILHDVYENKTYLVLEHVSCEGLFDYIVKKGRITP
      KEARPPROITISALDPCHSHSICHRODKPENNLLDEKNNIRIRIDFGMAISQVDSILE
      TSCGSPPIAPCEVITREKTYDKGRADWVSCVILFALKLVGLPDDNNLROLEKVRGG
      VFHHPHFITPDCOSLKGSEVDAARLLTLEHILVWYIGKNEEPPDIPRKVOI
      RSLPLESLDIDPDVLDVSMHSLGCFRRDKRLQDILSSEENDEKVIYFLLDKRERPSO
      EDLDPERNRINIDPRKRVSDPMLNRHGRKRRPERKSMSEVLTGGSGPVARRAITEAO
      HGRSRISISGASSGLSTPLSPSSRVPTPEPRRSPPTPTGEGVHNPPESGAPMPPT
      PPSPPSGVGVPMRARRLINSKSTLGSRRPFRRLKLOYTPPEMNLPESSPEAKSM
      FGNFTLSKEEEOIFVYIKDKPLSSIKADYHATLSLSPSSHSVTSQTSFAEKKATGG
      PAFFOKPVKROVDITYEGSGEAAKENGIIYVTITLISGSPRRKRVETITQAOQLSTH
      DPRAQQLSPPPPAPALSGMAGLGGKATVSTSSLS"
BASE COUNT      427 a      640 c      579 g      310 t
ORIGIN

```

Query Match 27.4%; Score 818; DB 9; Length 1956;
 Best Local Similarity 68.6%; Pred. No. 4,5e-144;
 Matches 1319; Conservative 0; Mismatches 410; Indels 195; Gaps 6;

Db	241	AGAAAGAACAAATCCCGCATTCGAGACTTTGGCATTTGGCCGTCCCTGCAGAGTTGGGACAGCC	300
Oy	642	TCCTGGAGACCAAGCTGGCGGTTCCCCCATTTATGCGTGTCCAGAGTGAATTAAAGGGAGAA	701
Db	301	TGTTGGAGACCAGCTGTGGGTCCCCCCCACTACGCCCTGCCCGAGGTGATCCGGGGAGAG	360
Oy	702	AATATGATGGCCGCCGGGACAGCATGTGGAGCGTGTGGAGTGCATCCATCTTGCCCTGTGCG	761
Db	361	AGTATGAGCGCGGAGGAGCGAGCGAGCTGTGGAGCTGGCGGCTCATCTCTGTTCGCTTGTG	420
Oy	762	TGGGGGCTCTGCCCCTTTGATGACGACAAACCTCCGCCAGCTGCTGAGAAAGTGAACGG	821
Db	421	TGGGGGCTCTGCCCCTTTCGACGATGACAACTTTCGACAGCTGCTGAGAAAGTGAACGG	480
Oy	822	GGGTCTTCACATGGCCCCACTTCATCTCCAGATTTGCCAGACCTCTCTGAGGGGATGA	881
Db	481	GGGTCTTCACATGGCCCCACTTCATCTCCAGATTTGCCAGACCTCTCTGAGGGGATGA	540
Oy	882	TGCAATGGAGCCCAAAAAAGGCTACGTGGAGCAAAATTCAGAAACATCTTGTGTACC	941
Db	541	GGAGATGTGAGCGCCGACGCGCGCTCAGCTAGAGACATTCAGAAACATATGTGTATA	600
Oy	942	TAGCGGGGAACACGAGCCAGACCCGTGCTTGGAGCCAGCCCTGGCGCCGGGTACCA	1001
Db	601	TAGGGGGCAAGAAATAGCCCGCAAC-----AGAGAGGCCCATCTTCGCAAGGTGAGA	654
Oy	1002	TGCGGAGCCGTGCATTCACAGGAGAGAGACTGGACCCCGCAGCTCTAGAGAGCATGATCAC	1061
Db	655	TCCGCTCGCTGCGCCACGCTTGGAGAGCATGACCCCGCAGCTGCTGGACAGCATGACTCAC	714
Oy	1062	TGCGCTGCTTCAGGAGCCGCGAGAGGCTGCATTCGCGAGCTGCGCAGTGAAGAGAGAAC	1121
Db	715	TGGGCTGCTTCAGGAGCCCAACAGCTCTGAGAGACCTGCTGCCGAGAGAGAGAAC	774
Oy	1122	AAGAAAGATGATATATTTATCTGCTTTTGGATGGGAAGGAGCGGTATCCAGCTGTGATG	1181
Db	775	AGGAAAGATGATTTACTTCTCTCTCTGTGACCGGAAAGAAAGTATACCCAGCAGAGG	834
Oy	1182	ACGAGAGCTGCTCCCGGAAATGATGTTGACCCGCCCGGAAAGCGTGTGATTTCTCCA	1241
Db	835	ATGAGAGACTGCCCCCGGAAAGATATGACCTTCGCCGGAAGCGTGTGACTCCCGA	894
Oy	1242	TGCTTACGCGTACGGGGAAGCGGCGACACAGAGGGGAATTCATGGAAGTCTGTAGCATCA	1301
Db	895	TGCTTAAACGGGACGGCAAGCGGCGGCGCAAGAAATTCATGGAAGTCTGTAGCATCA	954
Oy	1302	CCGATGCCCGGGGTGTGCTCCCTCTGTATCCACCGCAGCGGCGCTTGGATGTGCCAGC	1361
Db	955	C-----GGAGGGCGGCTCCCGGTGCTCGTCGCGCGCGGCGCATTTGAGATGGCCAGC	1005
Oy	1362	ACAGCCAGAGATCCCGTAGCGTCAGTGAGCGCTCCACGGGCTGTGTCTCCAGCCCTTAA	1421
Db	1006	ACGGCCAGAGGTCTCGGTCCATCAGGCTTCCTCAGGCTTTTCCACCAAGCCACCTCA	1065
Oy	1422	GGAGCCCAAGGAGTCCGGTCTTTCTCTTTTCAACGGAGCGGGGGCGTGGAGATGAGCTC	1481
Db	1066	GCAAGCCCGGGT-----	1078
Oy	1482	GAGGGGGGCTCCCGGACTTCCAAAAACGACGCTGCTTCTCGGGGCCCAAGGGGTG	1541
Db	1079	-----	1078
Oy	1542	GGGGCGCGGGGAGACGCCCCCGCCCGCAAGTCCGCTCACACCCCTGGCGGGCCCC	1601
Db	1079	-----GACCCCTACCCCTTCACCAAGGGGAGTCCCTCC-----	1115
Oy	1602	CAGGCTCCCGCGCTCTGTGGGGGAGCCCTTGTGACTCGCTTGCACAGCCGCCGGG	1661
Db	1116	-----ACCCCAAGGGGACACGTCTCCACAGCGCAAGG	1149
Oy	1662	CCAGTCCACCGGAGCCCGGGGACAAACACACCCCGCCAGCCCGCGGCTCGGGG	1721

Db	1150	AGAGCCCGGGTGGCCACGGCCCAACCCACGCCCCCGTCCAGGCCCC-----AGCGTCCGGAG	1203
Qy	1722	GAGCGCCCTGGAGAGATGCTGTCAATCTCCATTCGCAACAGATCTCTGGGCTCCCTCGCT	1781
Db	1204	GGGTGCTCTGGAGGGGCGGGGCTCAACTCCTCATCAGAAAGATCTTCTGGGCTCACCCGCT	1263
Qy	1782	TTTCACCGCGCGCAAGATGCGAGGTCCCTTACCGCGTGAGAGATGTCCAGCTTTGACGCGCAGT	1841
Db	1264	TCCACCGCGCGAAAACTGCAAGTTCGAGCGCCGGAGGAATGTCCAACTGTACACCAAGT	1323
Qy	1842	CCCTCCCGGAGCTGGCAAAAGCGCTCTGGTTCCGGAACTTCATCTCTTGGACAAAGAAG	1901
Db	1324	CGTCCCGAGAGCTGGCGAAGAAATCTCGTTTGGGAATTCATCAGCCTGGGAGAGAGAG	1383
Qy	1902	AACAAATTTCTCGTGGTAAAGGCAAACTCTAGCAGATCAAAACGATCATCGTCC	1961
Db	1384	AGCAATTTCTGTGGTGTATCAAAAGCAAACTCTAGTTCATCAAGGCTGTGACATCGTGC	1443
Qy	1962	ATGCGTTTCTGTGCATGCCCGCAGCTGAGTGCACAGTGTCTGTACACAGCACAGCTTCAGG	2021
Db	1444	ACGCTTCTCTGTGATTTCCCACTCTCAGCGACACAGGCTCATCTCCCAAAAGAGCTTCGGG	1503
Qy	2022	CCGAGTACAAAGGCCAGTGGCGGCCCTCGTTCGTTCCAAAAGCCGCTCCGCTTCCAAGTGG	2081
Db	1504	CCGAGTACAAAGGCCACGGGGGGGGCGACGCGTGTCCAGAAAGCCGSGTCAAGTTCACAGTTG	1563
Qy	2082	ACATCAGTCCCTCTAGAGGTCCAGAGCCCTCCCGCCGACGGGACGGGACGGAGGATGGTG	2141
Db	1564	AAATACACCTACACGAGAGGT-----GGGAGGCGCGCAGAAAGGAGAAACG	1605
Qy	2142	GCATCAGTCCGTCACCTTCACTCATCTCATCTGCGGGTCCAGCCGCTGGTTTCAAGCAGTGG	2201
Db	1606	GCATCTACTCCGTCACCTTCACTTCACTTCTGTCTAGAGGCCACGCGGTGCTTCAAGAGGGTGG	1665
Qy	2202	TGGAGACCATCCAGGCGACAGCTCTGAGACATCATGACCAGCCCTCCGTTGCAAGGCCCTGG	2261
Db	1666	TGGAGACCATCCAGGCGCCAGCTGCTGAGACACACAGACCCGCGCTCGGCGCCAGCACTTGT	1725
Qy	2262	CAGA 2265	
Db	1726	CAGA 1729	

RESULT 11
 AC020922
 LOCUS AC020922 134792 bp DNA linear PRI 01-AUG-2002
 DEFINITION Homo sapiens chromosome 19 clone CTD-2105E13, complete sequence.
 ACCESSION AC020922
 VERSION AC020922.8 GI:22038504
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 134792)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 134792)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 3 (bases 1 to 134792)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Submitted (01-AUG-2002) DOE Joint Genome Institute, 2800 Mitchell
 Drive, Walnut Creek, CA 94598, USA
 COMMENT On Aug 1, 2002 this sequence version replaced gi:15022017.
 Draft Sequence Produced by DOE Joint Genome Institute
 www.jgi.doe.gov
 Finishing Completed at Stanford Human Genome Center
 www-shgc.stanford.edu

OY 404 TACCTGGTCTGAGACGCTCTCGGGGGTGAGCTATTCGACTACTGTTAAAGAGGG 463
DB 62 TATTGATATTGGACATGATCCGGGAGAGCTCTTCGATTACGTGTGAAGAGGT 121
OY 464 AGACTGACGCCCAAGAGAGGCCGAAAGTTCTTCGCCAGATTGTCTGCGCTGACTTC 523
DB 122 CGATTGACGCCCAAGAGAGGCCGAAAGTTCTTCAGGCAATCATCTCCGCCCTGATTC 181
OY 524 TGCACAGCTACTCAATCTCCACAGAGACTTAAAGCCGGAACCTGCTTTGGATGAG 583
DB 182 TGCCTACTCGATTGATTCGATTCGCGACTTGAAGCCGGAATCTGCTGTGACGAG 241
OY 584 AAAACACATCCGATTCGAGACTTGGCATGGCGTCCCTGACAGTGGGGACAGCCTC 643
DB 242 AAGAAATTAATTAAGATAGCGAGCTTTGGATGCTTCCTGCAGCCAGCTGGCAGCATG 301
OY 644 CTGGAGACGCTGCGGGTCCGCCATTAATGCTGTCCAGAGGATTAAGGGGAAAA 703
DB 302 TTGGAGACCTCTGCGGCAAGCCCACTACGCTGTCCAGAGGTCATACGGGCGAGAG 361
OY 704 TATGATGCGCGCGGCGAGACATGTGAGCTGTGAGTCACTCTCGCCCTGCTG 763
DB 362 TACGATGGCGGCAAGCGGATGTCTGTCTGTGGGTCACTCTATGCTCTCTG 421
OY 764 GGGGCTCTGCTTGTATGACGAACTCTCGGAGCTGCTGGAAGAGTAAAGGGGC 823
DB 422 GGTGGCTTCCCTTCAGCAGCAACTTGCAGCTGCTGGAAGAAATCAAGCGGGC 481
OY 824 GTCTCCACATGCCCCACTTCATCTCCAGATTGCCAGACCTCTGAGGGGAATGTC 883
DB 482 GTCTTCACATACCGCATTTGTGCGCGGACTGCCAGAGTCTGTGCGGCGCATATT 541
OY 884 GAAGTGGACCCGAAAAAGGCTCAGTCTGAGCAAAATTCAGAAACATCTTGTACTTA 943
DB 542 GAGTCAATTCGGAGCGGGGCTCAGCTGGCTGAATCAACGGGCTCCGTGGGTACA 601
OY 944 GCGGGGAAACAGAGCCAAACCGTCTGAGAGCCAGCCCTGCGCGGAGTACCATG 1003
DB 602 GCTGCGCGCAAGGG--GAGCTGAGCTGAGCTGCCAATGATGAGTGTGACAGACA 658
OY 1004 CGGACCTGCAATCCAAAGAGAGTGAACCGGAGCTGCTAGAGCATGGCATCTG 1063
DB 659 CACGTATTCCCAAGACCCGCGGTGATCCGATGTGTTGAACGCGATTTGCTGCTG 718
OY 1064 GCGTCTTCAGGAGCGGAGAGGCTGCATCGGAGCTGGCAGTGGAGAGAGAACCAA 1123
DB 719 GCGTCTTCAGGAGAGAGAACTCATCCAGAGACTGCTCAGTTCAAGTCAATACG 778
OY 1124 GAAAAGATGATATTATTCCTTTGGATCGAAGAGAGCGGTATCCACGCTGTGAGAC 1183
DB 779 GAGAAGGTTATATTATTCCTGTTGTCGAGCGCAAGAGACGACCTGCGCTGAGAT 838
OY 1184 CAGGACCTGCTCCCGGAA 1203
DB 839 GATGATGAGATTGCGCAAAA 858

Search completed: April 21, 2003, 23:04:02
Job time : 5185.18 secs

XX 14-JUN-2001: 2001WO-US19444.
 PF 15-JUN-2000: 2000US-212073P.
 PR 23-JUN-2000: 2000US-213467P.
 PR 30-JUN-2000: 2000US-215651P.
 PR 07-JUL-2000: 2000US-216605P.
 PR 13-JUL-2000: 2000US-218372P.
 PR 25-AUG-2000: 2000US-228056P.
 XX (INCYTE GENOMICS INC.)
 PA Yue H, Lal P, Bandman O, Borowsky MU, Au-Young J, Lu Y;
 PI Gandhi AR, Tribouley CM, Walla NK, Yao MG, Lu DM, Greenwald SR;
 PI Ramkumar J, Griffin JA, Kearney L, Burford N, Nguyen DB, Tang YT;
 PI Baughn MR, He A, Thornton M, Hafalia A, Patterson C, Gururajan R;
 PI Lo TP, Khan F, Recipon SA, Azimzai Y, Policky JL, Ding L;
 PI Grether M, Elliott VS, Thangavelu K, Batra S, Iron CH;
 XX WPI: 2002-090307/12.
 DR P-PSDB: ABE16271.

PT New polypeptides, useful for diagnosing, treating or preventing
 PT disorders of growth and development, cardiovascular and lipid, and
 PT diseases such as cancer, comprise human kinase polypeptides

XX Claim 5; Page 188-189; 197pp; English.

XX The invention relates to human kinase PKIN proteins and their
 CC corresponding cDNAs. A composition containing PKIN agonist is useful for
 CC treating a disease or condition associated with decreased expression of
 CC PKIN and a composition comprising PKIN antagonist is useful for treating
 CC a disease or condition associated with overexpression of PKIN. The
 CC disorders include cancer (leukemia, adenocarcinoma, lymphoma, melanoma,
 CC myeloma, sarcoma, teratocarcinoma, Hodgkin's disease); immune disorder
 CC (Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease,
 CC atherosclerosis, anemia, allergies, adult respiratory distress syndrome,
 CC autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes
 CC mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease,
 CC osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome,
 CC rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis,
 CC bacterial, parasitic, fungal, viral, protozoal and helminthic infections)
 CC growth and development disorders (arteriosclerosis, cirrhosis, hepatitis,
 CC Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts); cardio
 CC vascular disease (arteriovenous fistula, hypertension, vasculitis,
 CC aneurysms, congestive heart failure, angina pectoris, myocarditis,
 CC ischemic heart disease, chronic bronchitis, lung tumours); lipid
 CC disorder (fatty liver, Fabry's disease, Niemann-Pick's disease,
 CC hypochlolesterolemia, obesity). PKIN DNA is useful for assessing
 CC toxicity of a test compound and in gene therapy. The present sequence
 CC is human PKIN-17 cDNA.

XX Sequence 2897 BP: 599 A; 921 C; 877 G; 500 T; 0 other;

XX Query Match 91.3%; Score 2722.2; DB 24; Length 2897;
 XX Best Local Similarity 99.1%; Pred. No. 0;
 XX Matches 2747; Conservative 0; Mismatches 23; Indels 1; Gaps 1;

DB 127 CACGCCCAATATGTGGGCCCTATCGGCTGAGAAAGACGCTGGCAAGAGACAGACAGG 186
 QY 164 CACGCCCAATATGTGGGCCCTATCGGCTGAGAAAGACGCTGGCAAGAGACAGACAGG 223
 DB 187 CTGGTTAAACCTGGGGCTGCATGCATCAGGGGTCAAGAGGTCCGCCATCAAGATCGTGAAC 246
 QY 284 CGGAGAAAGCTGTGGAGTGGTGTGATGAGAGTGAGAGCGGAGATCGCATCTGAAG 343
 DB 247 CGGGAGAGAGCTGTGGAGTGGTGTGATGAGAGTGAGAGCGGAGATCGCATCTGAAG 306
 QY 344 CTCATCGAAGCCCAAGATGCTCAAGCTCAGACGAGCTTACGAGAAAGAAATATTTG 403
 DB 307 CTCATCGAAGCCCAAGATGCTCAAGCTCAGACGAGCTTACGAGAAAGAAATATTTG 366

QY 404 TACCTGTTCTTGAGACAGCTCTGGGGGGTGGAGCTATTGACTACCTGTTAAAGAGGG 463
 DB 367 TACCTGTTCTTGAGACAGCTCTGGGGGGTGGAGCTATTGACTACCTGTTAAAGAGGG 426
 QY 464 AGACTGAGCCCAAGAGAGGCCCAAGATTCTTCGCCAGATTGTTGTGGCTGGACTTC 523
 DB 427 AGACTGAGCCCAAGAGAGGCCCAAGATTCTTCGCCAGATTGTTGTGGCTGGACTTC 486
 QY 524 TGGCAGAGTACTCCATCTGCGCAGAGACTTAAAGCCCAAGACCTGCTTTGGATGAG 583
 DB 487 TGGCAGAGTACTCCATCTGCGCAGAGACTTAAAGCCCAAGACCTGCTTTGGATGAG 546
 QY 584 AAAAACCAATCCGATCGATTCAGACTTCGCGATGCGCTCCCTGCAAGTGGGGAGACGCTC 643
 DB 547 AAAAACCAATCCGATCGATTCAGACTTCGCGATGCGCTCCCTGCAAGTGGGGAGACGCTC 606
 QY 644 CTGAGAGACAGCTGGCGGGTCCCAATATGCTGTCAGAGAGGTGATTAAAGGGGAAAAA 703
 DB 607 CTGAGAGACAGCTGGCGGGTCCCAATATGCTGTCAGAGAGGTGATTAAAGGGGAAAAA 666
 QY 704 TATGATGGCCGGCGGAGACAGATGTGAGCTGTGAGTATCTCTTCCGCTGCTGTCG 763
 DB 667 TATGATGGCCGGCGGAGACAGATGTGAGCTGTGAGTATCTCTTCCGCTGCTGTCG 726
 QY 764 GGGGCTTGCCCTTTGATGAGAGCAACTCCGACAGCTGCTGAGAGAGTGAAGGGG 823
 DB 727 GGGGCTTGCCCTTTGATGAGAGCAACTCCGACAGCTGCTGAGAGAGTGAAGGGG 786
 QY 824 GCTTTCACATGCCCCACATCTTCTCCAGATGTCCAGAGCCTCTGAGGGAAATGATC 883
 DB 787 GCTTTCACATGCCCCACATCTTCTCCAGATGTCCAGAGCCTCTGAGGGAAATGATC 846
 QY 884 GAAGTGAAGCCCGAAAAAGGCTCACTGTGAGCAATTCAGAAACATCTTGTTACTCA 943
 DB 847 GAAGTGAAGCCCGAAAAAGGCTCACTGTGAGCAATTCAGAAACATCTTGTTACTCA 906
 QY 944 GGGGGGAAACACAGACAGACCCGTCGAGAGCCAGCCCTGCGCCCGGGTGAACATG 1003
 DB 907 GGGGGGAAACACAGACAGACCCGTCGAGAGCCAGCCCTGCGCCCGGGTGAACATG 966
 QY 1004 GGGAGCTGCAATCCAGAGAGAGTGGAGACCCCGAGCTCTAGAGAGCATGCTACTG 1063
 DB 967 GGGAGCTGCAATCCAGAGAGAGTGGAGACCCCGAGCTCTAGAGAGCATGCTACTG 1026
 QY 1064 GGGTCTTCAGGAGACCGGAGAGCGTCATCGCGAGCTGCGAGTGAAGAGGAGAAACAA 1123
 DB 1027 GGGTCTTCAGGAGACCGGAGAGCGTCATCGCGAGCTGCGAGTGAAGAGGAGAAACAA 1086
 QY 1124 GAAAAATGATATATATCTGCTTTGGATCGGAAGAGAGCGGATATCCAGCTGTAGAGAC 1183
 DB 1087 GAAAAATGATATATATATCTGCTTTGGATCGGAAGAGAGCGGATATCCAGCTGTAGAGAC 1146
 QY 1184 CAGGACCTGCTCCCGCGGAATATGTTGACCCCGCGGAAGCGTGTGATTTCCCATG 1243
 DB 1147 CAGGACCTGCTCCCGCGGAATATGTTGACCCCGCGGAAGCGTGTGATTTCCCATG 1206
 QY 1244 CTGAGCCCTCAGGAGAACCGGAGACAGAGCGGAAGTGCATGGAAGTCTTGAGCATCAC 1303
 DB 1207 CTGAGCCCTCAGGAGAACCGGAGACAGAGCGGAAGTGCATGGAAGTCTTGAGCATCAC 1266
 QY 1304 GATGCGGGGGTGGTGGCTCCCTGTACCCACCGAGGGGCTTGGAGATGGCCAGAC 1363
 DB 1267 GATGCGGGGGTGGTGGCTCCCTGTACCCACCGAGGGGCTTGGAGATGGCCAGAC 1326
 QY 1364 AGCCAGAGATCCCGTAAAGTGAAGTGAAGCGTCCAGAGGCTCTCCAGCCCTGAAC 1423
 DB 1327 AGCCAGAGATCCCGTAAAGTGAAGTGAAGCGTCCAGAGGCTCTCCAGCCCTGAAC 1386
 QY 1424 AGCCAGAGATCCCGTCTTTCTTTTCCATTCACGAGAGCGGGGGCTGAGATAGGCTGCA 1483
 DB 1387 AGCCAGAGATCCCGTCTTTCTTTTCCATTCACGAGAGCGGGGGCTGAGATAGGCTGCA 1446

OY	1484	GGCGGGGGCTCCCGCACTTCCAAAAGCAGAGCTCTTCTCGGGGCCCCAGGGGTGGG	1543
Db	1447	GGCGGGGGGGCTCCCGCACTTCCAAAAGCAGAGAGGTCTTCTCGGGGCCCCAGGGGTGGG	1506
OY	1544	GGCGCGGGGGAGCAGCCCCCGCGCCAGTGGCCCGCTGCACACCCCTGCCGGCCCCCA	1603
Db	1507	GGCGCGGGGGAGCAGCCCCCGCGCCCAATGGCCCGCTGCACACCCCTGCCGGCCCCCA	1566
OY	1604	GGCTCCCCCGCGCTCTCTCGGGGGGAGCCCCCTTGCATCGCCCTCTGACAGAGCCCCGGGGC	1663
Db	1567	GGCTCCCCCGCGCTCTCTCGGGGGGAGCCCCCTTGCATCGCCCTCTGACAGAGCCCCGGGGC	1626
OY	1664	AGTCCACCGGGAGCCCGGGGGCAACACCACCCCGCAGCCCGGGCGGTGGCGTCCGGGGGA	1723
Db	1627	AGTCCACCGGGAGCCCGGGGGCAACACCACCCCGCAGCCCGGGCGGTGGCGTCCGGGGGA	1686
OY	1724	GGCGCGCTGGAGAGTCTCTCAACTTCATCCCGCAACAGCTTCTCGGGGTCTCCCTCGCTTT	1783
Db	1687	GGCGCGCTGGAGAGTCTCTCAACTTCATCCCGCAACAGCTTCTCGGGGTCTCCCTCGCTTT	1746
OY	1784	CACCGCGCAAGATCAGGTCTCCCTACCGCTGAGAGATGTCCAGTCTGAGCGCAGAGTCC	1843
Db	1747	CACCGCGCAAGATCAGGTCTCCCTACCGCTGAGAGATGTCCAGTCTGAGCGCAGAGTCC	1806
OY	1844	TCGCCGGAGCTGGCAAAACGCTCTCTGTTCGGGACTTCATCTCTTGGCAAAAGAA	1903
Db	1807	TCGCCGGAGCTGGCAAAACGCTCTCTGTTCGGGACTTCATCTCTTGGCAAAAGAA	1866
OY	1904	CAAAATTTCTCTGTGTAAAGGCAAAACCTTCAGAGAGCTCAAGAGCAGCATCTGTCAT	1963
Db	1867	CAAAATTTCTCTGTGTAAAGGCAAAACCTTCAGAGAGCTCAAGAGCAGCATCTGTCAT	1926
OY	1964	GCCTTCTGTCTCATCCCGCAGCCTGAGTCAACAGTGTCTGTACAGACACAGCTTCAGGGCC	2023
Db	1927	GCCTTCTGTCTCATCCCGCAGCCTGAGTCAACAGTGTCTGTACAGACACAGCTTCAGGGCC	1986
OY	2024	GAGTACAAAGGCCAGTGGCGGGCCCTCCGTCTTCCAAAAGCCGCTCCGCTTCCAGAGTGGAC	2083
Db	1987	GAGTACAAAGGCCAGTGGCGGGCCCTCCGTCTTCCAAAAGCCGCTCCGCTTCCAGAGTGGAC	2046
OY	2084	ATCAGCTCCCTCTGAGAGGTTCTCAGAGCCCTCCCGCCGACGAGCAGCGAGCGAGTGTGGC	2143
Db	2047	ATCAGCTCCCTCTGAGAGGTTCTCAGAGCCCTCCCGCCGACGAGCAGCGAGCGAGTGTGGC	2106
OY	2144	ATCTACTCCGTCACTTCACTCTCATCTGGGTCCCAAGCCGTGGTTCAAGCGAGTGGTG	2203
Db	2107	ATCTACTCCGTCACTTCACTCTCATCTGGGTCCCAAGCCGTGGTTCAAGCGAGTGGTG	2166
OY	2204	GAGACCATCAAGCAGACGTCTCTGAGACACTCATGACACAGCCCTCCGCGAGGCCCTGGCA	2263
Db	2167	GAGACCATCAAGCAGACGTCTCTGAGACACTCATGACACAGCCCTCTCGTCAGGCCCTGGCA	2226
OY	2264	GACGAGAAAGACGGGGGCCAGACCCCGGCTGTGGTGGCCCCACCCCGAAGCCTTGACACC	2323
Db	2227	GACGAGAAAGACGGGGGCCAGACCCCGGCTGTGGTGGCCCCACCCCGAAGCCTTGACACC	2286
OY	2324	GCACCCGGCGGCCACAGACCCAGAGCTGAGAGTCTCCCGCGGAGGCCCCCCCAAGAGC	2383
Db	2287	GCACCCGGCGGCCACAGAGCTGAGAGTCTCTCCCGCGGAGGCCCCCCCAAGAGC	2346
OY	2384	AAGAAAGCTCTGTGGCCACAAAGGGAGCCCTCTCCCTGACCCCAAGGGGGCGGGAGGGA	2443
Db	2347	AAGAAAGCTCTGTGGCCACAAAGGGAGCCCTCTCCCTGACCCCAAGGGGGCGGGAGGGA	2406
OY	2444	GGGGAGCCCCCTTCACCCCTTCTCGGTGGCCCCCACTGTGAATCTTAAATAGGCCCA	2503
Db	2407	GGGGAGCCCCCTTCACCCCTTCTCGGTGGCCCCCACTGTGAATCTTAAATAGGCCCA	2466
OY	2504	AGGAACATGTCCGGAGGGGGGTGGAACAATAAAACCGGCTTGGCCCTCAGGGAGTGGCGT	2563
Db	2467	AGGAACATGTCCGGAGGGGGGTGGAACAATAAAACCGGCTTGGCCCTCAGGGAGTGGCGT	2526
OY	2564	CCACAGCGCCGTCCCAATGGGGGTGGTTCTAGGGACAAAGGGGGCGGGGAGCTGTTTTC	2623

Dd	2527	CCACGAGGC CGATGCCAACTGGCGGGTGGTTCTTAGGGGAACAAGGGGGCGGGGAGCTGTTC	2586
Oy	2634	TATTATTATTATGATTGATTTAATTATTATTATTATTATTATTATGATCAATCTCTCTGGGGGTGGG	2693
Dd	2587	TATTATTATTATGATWTATTTAATTATTATTATTATTATTATGATCAATCTCTCTGGGGGTGGC	2646
Oy	2684	GTCGGGGAGGACGACGAGACTGGTTGGGGTGAGCTTAGCATCCGGACAGGGGCCCTGTGTC	2743
Dd	2647	GTGGGGGAGGACGACGAGACTGGTTGGGGTGAGCTTAGCATCCGGACAGGGGCCCTGTGTC	2706
Oy	2744	CCTGTGTGTGCCCAACCCCCTCTCTCCGGGCCCTCTCTCCCTGTGTC -TCCCCCACG	2802
Dd	2707	CCTGTGTGTGCCCAACCCCCTCTCTCCGGGCCCTCTCTCCCTGTGTC -TCCCCCACG	2766
Oy	2803	ACCTCTCTGAGGATTTGCTCTCCGGAGAAGATTTCTGTTTCGGGTGATTCGCGCTGGCT	2862
Dd	2767	ACCTCTCTGAGGATTTGCTCTCCGGAGAAGATTTCTGATTAACGGGTGATCTGCTGGCT	2826
Oy	2863	CCGTGTCTCTGATTTCCGGCGGCGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATA	2922
Dd	2827	CCGTGTCTCTGATTTCCGGCGGCGGCAAAAAAAAAAACACACACCAACACAAAGGGCAC	2886
Oy	2923	AATAAATAAA 2933	
Dd	2887	AACAAAAAAA 2897	
RESULT 2			
AAS06717			
ID	AAS06717 standard; cDNA: 2385 BP.		
XX	AAS06717;		
AC			
XX			
DT	12-SEP-2001 (first entry)		
DE	Polynucleotide sequence encoding human protein kinase #17.		
KM	Human; protein kinase; PK; STK; cancer; cardiovascular disease;		
KW	metabolic disorder; immune related disease; neurological disorder;		
KN	neurodegenerative disorder; inflammatory disorder; infectious disease;		
XX	reproductive disorder; gene therapy; ss.		
OS	Homo sapiens.		
XX			
PN	WO200138503-A2.		
XX			
PD	31-MAY-2001.		
XX			
PF	22-NOV-2000; 2000MO-US32085.		
XX			
PR	24-NOV-1999; 99US-0167482.		
XX			
PA	(SUGEN-) SUGEN INC.		
XX			
PI	Plozman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;		
XX	Flanagan P, Clary D;		
DR	WPI; 2001-343950/36.		
XX	P-PSDB; AAU03517.		
PT	Nucleic acids encoding human kinase polyptides, useful for preventing		
PT	diagnosing and/or treating e.g. cancer, immune, cardiovascular and		
XX	neuronal-associated diseases, and microbial infections -		
XX			
PS	Example 1; Figure 1; 433pp; English.		
XX			
CC	AAS06701-AAS06757 encode for novel human protein kinases #1-57. The		
CC	novel protein kinases have been identified as members of the tyrosine		
CC	or serine/threonine kinase (PK and STK) families. The polynucleotides		
CC	encoding protein kinases and the polypeptides may be used in the		
CC	prevention, diagnosis and treatment of diseases associated with		
CC	inappropriate kinase expression. For example, they may be used to treat		

CC cancers (especially cancers of hematopoietic origin), cardiovascular
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
 CC immune related diseases (e.g. rheumatoid arthritis), neurological
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
 CC Additionally, polynucleotides encoding protein kinases may be
 CC used for gene therapy and as DNA probes in diagnostic assays.
 CC The protein kinase polypeptides may be used as antigens in the production
 CC of antibodies against the protein kinases and in assays to identify
 CC modulators of protein kinase expression and activity.
 CC
 XX
 SQ Sequence 2385 BP; 494 A; 768 C; 731 G; 392 T; 0 other;

Query Match 75.8%; Score 2259; DB 22; Length 2385;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 164 CAGGCCAATATGTTGGGCTTATGCTGCTGAGAGAGAGCTGGGCAAGGACAGACAGG 223
 DB 127 CAGGCCAATATGTTGGGCTTATGCTGCTGAGAGAGAGCTGGGCAAGGACAGACAGG 186
 OY 224 CTGTTAACTCGGGGCTCAGTCTCAGTCTCAGAGGCTCGCCATCAAGATCTGAC 283
 DB 187 CTGTTAACTCGGGGCTCAGTCTCAGTCTCAGAGGCTCGCCATCAAGATCTGAC 246
 OY 284 CGGAGAGCTGTGCGAGTGTGCTGTGATGAGGTGAGGCGGAGATGCCATCTGAAG 343
 DB 247 CGGAGAGCTGTGCGAGTGTGCTGTGATGAGGTGAGGCGGAGATGCCATCTGAAG 306
 OY 344 CTCATGACACCCATGCTCTCAAGCTCCAGAGCTACAGAGCTACAGAGCAAGAAATTTG 403
 DB 307 CTCATGACACCCATGCTCTCAAGCTCCAGAGCTACAGAGCTACAGAGCAAGAAATTTG 366
 OY 404 TACGTGTTCTGAGAGCAGTCTGCGGGGTGAGCTATTCTACTACTCTGTAAGAAAGGG 463
 DB 367 TACGTGTTCTGAGAGCAGTCTGCGGGGTGAGCTATTCTACTACTCTGTAAGAAAGGG 426
 OY 464 AGAGTACGACCCCAAGAGAGGCGGAAAGTTCTTCCGCAAGTTGTGTGCTGCTGAGCTTC 523
 DB 427 AGAGTACGACCCCAAGAGAGGCGGAAAGTTCTTCCGCAAGTTGTGTGCTGCTGAGCTTC 486
 OY 524 TCCACAGCTACTCTCATCTGCGCAGAGAGCTTAAGGCGGCAAGCTTTGGATGAG 583
 DB 487 TCCACAGCTACTCTCATCTGCGCAGAGAGCTTAAGGCGGCAAGCTTTGGATGAG 546
 OY 584 AAAAACAACATCCGATTCAGATTCGCGATGCGCTTCCGAGGTGGGGAGAGCCTC 643
 DB 547 AAAAACAACATCCGATTCAGATTCGCGATGCGCTTCCGAGGTGGGGAGAGCCTC 606
 OY 644 CTGGAGACAGCTGGGGGCTCCCTCATATGCTGTCCAGAGGTGAATTAAGGGGAAAA 703
 DB 607 CTGGAGACAGCTGGGGGCTCCCTCATATGCTGTCCAGAGGTGAATTAAGGGGAAAA 666
 OY 704 TATGATGGCGCGCGGCAAGATGTGAGCTGTGAGTCACTCTTCCGCTGCTGCTG 763
 DB 667 TATGATGGCGCGCGGCAAGATGTGAGCTGTGAGTCACTCTTCCGCTGCTGCTG 726
 OY 764 GGGGCTGCTGCTTGTGATGAGCAACAACCTCCGCACTGCTGAGAGAGGTGAACCGGGC 823
 DB 727 GGGGCTGCTGCTTGTGATGAGCAACAACCTCCGCACTGCTGAGAGAGGTGAACCGGGC 786
 OY 824 GTCCTTCCATGCTCCCTCATCTTCTCCAGATGTGCAAGAGCTCTGAGGGAGATGATC 883
 DB 787 GTCCTTCCATGCTCCCTCATCTTCTCCAGATGTGCAAGAGCTCTGAGGGAGATGATC 846
 OY 884 GAATGAGAGCCGAAAGAGCTCAGTCTGAGCAAAATTCAGAAACATCTTGGTACTCA 943
 DB 847 GAATGAGAGCCGAAAGAGCTCAGTCTGAGCAAAATTCAGAAACATCTTGGTACTCA 906
 OY 944 GCGGAGAAACAGAGCCAGAGCCGTGCTGAGAGCCAGCCCTGCGCGCGGTAGCCATG 1003
 DB 907 GCGGAGAAACAGAGCCAGAGCCGTGCTGAGAGCCAGCCCTGCGCGCGGTAGCCATG 966

OY 1004 CGAGAGCTGCTCCATCCAAAGAGAGCTGAGACCCGAGCTCTTAAGAGACATGACACTG 1063
 DB 967 CGAGAGCTGCTCCATCCAAAGAGAGCTGAGACCCGAGCTCTTAAGAGACATGACACTG 1026
 OY 1064 GCGTCTTAAAGAGAGCCGAGAGGCTGATCGGAGCTGCGCAGTGAAGAGAGAACAA 1123
 DB 1027 GCGTCTTAAAGAGAGCCGAGAGGCTGATCGGAGCTGCGCAGTGAAGAGAGAACAA 1086
 OY 1124 GAAAGATATATATATATGCTTTGATGAGAGAGAGGAGATCCACTGTGAGAGAC 1183
 DB 1087 GAAAGATATATATATATGCTTTGATGAGAGAGAGGAGATCCACTGTGAGAGAC 1146
 OY 1184 CAGAGAGCTGCTCCCGGAGATGATGTTGACCCCGGAGAGGCTGATGCTCCATG 1243
 DB 1147 CAGAGAGCTGCTCCCGGAGATGATGTTGACCCCGGAGAGGCTGATGCTCCATG 1206
 OY 1244 CTGAGCCGTCACGAGAGAGCGGAGCCAGAGAGGAGTCCATGGAAGTCTGAGCATCAC 1303
 DB 1207 CTGAGCCGTCACGAGAGAGCGGAGCCAGAGAGGAGTCCATGGAAGTCTGAGCATCAC 1266
 OY 1304 GATGCGGGGGGAGTGGGCTCCCTGTACCAACCCGAGCGGCTTGGAGATGCGCCAGAC 1363
 DB 1267 GATGCGGGGGGAGTGGGCTCCCTGTACCAACCCGAGCGGCTTGGAGATGCGCCAGAC 1326
 OY 1364 AGCCAGAGATCCCGTAGAGCTCAGTGAAGCTCCACGAGTCTGCTCCAGCCCTTAAGC 1423
 DB 1327 AGCCAGAGATCCCGTAGAGCTCAGTGAAGCTCCACGAGTCTGCTCCAGCCCTTAAGC 1386
 OY 1424 AGCCCAAGAGAGTCCGCTCTTTTCCACCGAGAGCCGAGGCTGAGATGAGCTCGA 1483
 DB 1387 AGCCCAAGAGAGTCCGCTCTTTTCCACCGAGAGCCGAGGCTGAGATGAGCTCGA 1446
 OY 1484 GGGGGGGGCTCCCGCACTTCCAAAGAGAGAGGCTCTTCTGGGGGCGCCAGGGGTGGG 1543
 DB 1447 GGGGGGGGCTCCCGCACTTCCAAAGAGAGAGGCTCTTCTGGGGGCGCCAGGGGTGGG 1506
 OY 1544 GGGGCGGGGAGAGAGCCCGGCGGCTCCAGTGGCGGCTCCAGACCCCTGCGGCGCCCA 1603
 DB 1507 GGGGCGGGGAGAGAGCCCGGCGGCTCCAGTGGCGGCTCCAGACCCCTGCGGCGCCCA 1566
 OY 1604 GGGTCCCGGCGCTCTCTGAGGAGAGCCCGGCTGAGCTGCTCTGCAACAGCCCGGGCC 1663
 DB 1567 GGGTCCCGGCGCTCTCTGAGGAGAGCCCGGCTGAGCTGCTCTGCAACAGCCCGGGCC 1626
 OY 1664 AGTCCACAGAGAGAGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1723
 DB 1627 AGTCCACAGAGAGAGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1686
 OY 1724 GCGGCTGAGAGAGTCTCTCAATCCATCGCAACAGCTTCTGAGGCTCCCTCGCTTT 1783
 DB 1687 GCGGCTGAGAGAGTCTCTCAATCCATCGCAACAGCTTCTGAGGCTCCCTCGCTTT 1746
 OY 1784 CAGCGGGGCAAGATGAG 1843
 DB 1747 CAGCGGGGCAAGATGAG 1806
 OY 1844 TCCCGGAGAGTGGCAAAAGAGCTGCTGAGGAGCTTCCATCTTGGAGAAAGAGAA 1903
 DB 1807 TCCCGGAGAGTGGCAAAAGAGCTGCTGAGGAGCTTCCATCTTGGAGAAAGAGAA 1866
 OY 1904 CAAATATCTCTGCTGCTAAAGAGCAAACTCTCAGAGAGATCAAAAGCAAGATCTGAT 1963
 DB 1867 CAAATATCTCTGCTGCTAAAGAGCAAACTCTCAGAGAGATCAAAAGCAAGATCTGAT 1926
 OY 1964 GCGTCTGAGAGATCCAGAGCTGAGTCAAGTGTGCTGCAAGAGAGCTTCAAGGCGC 2023
 DB 1927 GCGTCTGAGAGATCCAGAGCTGAGTCAAGTGTGCTGCAAGAGAGCTTCAAGGCGC 1986
 OY 2024 GAGTACAAG 2083
 DB 1987 GAGTACAAG 2046

[illegible]

Db 1702 TGCAGCCTTCCTGATTCAGATCCAGTCTCAGCCAGCAGCTCATCTCCCAAGAGCTTCC 1761
 QY 2019 GGGCGAGTACAAAGCCAGTGGCGGCCCTTCGCTTTCAAAAGCCGCTCCGCTTCAGG 2078
 Db 1762 GGGCCAGATCAAGGGCCAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1821
 QY 2079 TGGACATCAGCTCCTGCTGAGGCTCCAGAGCCCTCCCGGAGCGGAGCGGAGGTG 2138
 Db 1822 TTGATATTCATCCTACAGGAGGGT-----GGGGAAGCGCCAGAGAGAGA 1863
 QY 2139 GTGGCATCTACTCCCTCACCCTTCTCATCTCGGGTCCAGCCGCTGCTTCAAGCGAG 2198
 Db 1864 AGGGCATCTACTCCCTCACCCTTCTCATCTCGGGTCCAGCCGCTGCTTCAAGAGGG 1923
 QY 2199 TGGTGAGACATCCTCAGGAGCAGCTCCTGAGCACTGACAGAGCCCTCCGTCAGAGCC 2258
 Db 1924 TGGTGAGACATCCTCAGGAGCAGCTCCTGAGCACTGACAGAGCCCTCCGTCAGAGCT 1983
 QY 2259 TGGCAGA 2265
 Db 1984 TGTTCAGA 1990

RESULT 5

AD34315
 ID AAD34315 standard; cDNA; 2647 BP.

AD34315;

16-JUL-2002 (first entry)

Human PKIN-18 cDNA.

Human; kinase; enzyme; PKIN-18 protein; immune system disorder; anaemia; acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease; asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease; AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma; leukemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder; Down's syndrome; gene therapy; protein therapy; cytosolic; gene; ss.

Homo sapiens.

Key Location/Qualifiers
 1.1995
 CDS /product="Human PKIN-18 protein"

MO200218557-A2.

07-MAR-2002.

31-AUG-2001; 2001WO-US27219.

31-AUG-2000; 2000US-229873P.

08-SEP-2000; 2000US-231357P.

14-SEP-2000; 2000US-232654P.

22-SEP-2000; 2000US-234902P.

29-SEP-2000; 2000US-236495P.

06-OCT-2000; 2000US-238389P.

13-OCT-2000; 2000US-240542P.

(INCY-) INCYTE GENOMICS INC.

Bandman O, Nguyen DB, Walla NK, Hafalia AFA, Yao MG, Gandhi AR;

PI Gururajan R, Ding L, Patterson C, Yue H, Baughn MR, Tribouley CM;

PI Thurnston M, Elliott VS, Lu Y, Ison CH, Au-Yang J, Tang YT;

PI Azimzai Y, Burdill JD, Marcus GA, Zingler KA, Lu DM, Lai PG;

PI Rankumar J, Warren BA, Kearney L, Policky JL, Thangavelu K;

PI Burford N;

XX MPI; 2002-329769/36.

DR P-PSDB; AAE21723.

XX

PT New human kinases, useful for diagnosing, treating or preventing immune
 PT system disorders (e.g. Crohn's disease), neurological disorders (e.g.
 PT epilepsy), or cell proliferative disorders (e.g. cancers such as
 PT leukemia or lymphoma) -
 PS Claim 97; Page 212-213; 218pp; English.

CC The present invention relates to human kinases (PKIN) and polynucleotides
 CC encoding such proteins. PKIN sequences of the invention are useful for
 CC diagnosing, treating or preventing disorders associated with aberrant
 CC expression of PKIN, particularly immune system disorders (e.g. acquired
 CC immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,
 CC anaemia, asthma), neurological disorders (e.g. epilepsy, Charcot-Marie-
 CC Tooth disease or seizures), cell proliferative disorders (e.g. cancers
 CC such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma),
 CC and developmental disorders (e.g. Down's syndrome). They are also used
 CC in gene therapy and protein therapy. The present sequence is a cDNA
 CC encoding human PKIN-18 protein.

XX Sequence 2647 BP; 525 A; 885 C; 789 G; 448 T; 0 other;

Query Match 32.1%; Score 955.8; DB 24; Length 2647;

Best Local Similarity 70.0%; Pred. No. 1.2e-177;

Matches 1473; Conservative 0; Mismatches 437; Indels 195; Gaps 6;

QY 173 TATGTGGGCCCCCTATCGGCTGAGAGAGACGCTGGCAAGAGACAGAGGCTGTTAA 232
 Db 31 TATGTGGGCCCCCTATCGGCTGAGAGAGACGCTGGCAAGAGAGAGAGGCTGTTAA 90
 QY 233 CTGGGGGTCATCATCATCAGGGGTCGAGAGGTCGCCATCAAGTCTGAAACCGGAGAG 232
 Db 91 CTGGGGGTCATCATCATCAGGGGTCGAGAGGTCGCCATCAAGTCTGAAACCGTGAAG 150
 QY 293 CTGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 352
 Db 131 CTGACGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 210
 QY 353 CACCCACATGCTCTCAAGCTCCAGCAGAGCTCTAGAGAGAGAGAGAGAGAGAGAGAG 412
 Db 211 CACCCACATGCTCTCAAGCTCCAGCAGAGCTCTAGAGAGAGAGAGAGAGAGAGAGAG 270
 QY 413 CTGGAGCAGCTCTCGGGGGGTCGAGATTCGACTACTGCTGTAAGAGAGAGAGAGAG 472
 Db 271 CTAGAGCAGCTCTCGGGGGGTCGAGATTCGACTACTGCTGTAAGAGAGAGAGAGAG 330
 QY 473 CCAAG 532
 Db 331 CCTAAG 390
 QY 533 TATCTCATCTGCTCCAG 592
 Db 391 CACTTCATATGCTCCAG 450
 QY 593 ATCCGATTTGAG 652
 Db 451 ATCCGATTTGAG 510
 QY 653 AGCTGCGGGTCCCGCCATATGCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 712
 Db 511 AGCTGCGGGTCCCGCCATATGCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 570
 QY 713 CGCGGGGAG 772
 Db 571 CGGAAGGGGAG 630
 QY 773 CCTTTGATGAG 832
 Db 631 CCTTTGATGAG 690
 QY 833 ATGCCCACTTCACTCTCCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 892
 Db 691 ATGCCCACTTCACTCTCCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 750

OY 893 CCGGAAAAAGGCTACGTGAGCAAAATTCGTTGTTACTTACCTAGGCGGAAA 952
 DB 751 GCCGACAGCCGCTCAGCTGAGCAGATTCAGAAACACATATGATATAGGGGCAAG 810
 OY 953 CAGCAGCAGACCCGCTGAGCAGCAGCCCTGGCCGCGGAGTATGCGAGCACTG 1012
 DB 811 AATGAGCCGCAAGC-----AGAGAGCCATCTCTCGAAGGTGCAGATCCGCTGCTG 864
 OY 1013 CCATGCAAGAGAGCTGAGCCGAGCTCTAGAGAGCAGGATCACTGAGGCTGCTTC 1072
 DB 865 CCGAGCTGAGAGCAGATCAGACCCGACCTGCTGACACATCAGTCACTGAGCTGCTTC 924
 OY 1073 AGGAGCCGCGAGAGAGCTGAGAGTGGGAGTGGGAGTGGAGAGAGCAACCAAGAAAGATG 1132
 DB 925 CGAGACCGCAACAGCTGCTCAGAGACCTGCTGCTCCGAGAGAGAGAAACAGAGAAATG 984
 OY 1133 ATATATATATCTGCTTTTGTATGAGAGAGAGAGGTATCCAGCTGTGAGAGCAGCACTG 1192
 DB 985 ATTTACTTCCCTCTGCTGAGCAGGAAAGAAAGTACCCGAGCAGAGAGATGAGACCTG 1044
 OY 1193 CCTCCCGGAAATGATGTTGACCCCGGAGAGCGTGTGATTTCTCCATGCTGAGCCGT 1252
 DB 1045 CCCCCCGGAAAGAGATAGACCTCCCGGAAAGCTGTGAGATCTCCCGATGCTGAACCGG 1104
 OY 1253 CAGGGAAGGCGGAGAGAGAGAGTCAAGTCAAGTCTTCAAGATCAACGATGCGGAG 1312
 DB 1105 CAGGCAAGCGGCGGCGGAG 1155
 OY 1313 GGTGAGGCTCTCCCTGACCCAGCAGGAGCTGTGAGAGTGGCCAGAGAGAGAGAGAG 1372
 DB 1156 GAGGCGGCTCTCCGAGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1215
 OY 1373 TCCGCTAGCTGAGTGAAGCTTCAAGGAGCTGTCTCAGCCCTCAAGAGAGAGAGAG 1432
 DB 1216 TCTGCTGCTACAG 1270
 OY 1433 AGTCCGCTCTTTCTTTTACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1492
 DB 1271 ----- 1270
 OY 1493 TCCCGAGCTTCAAAAG 1552
 DB 1271 -----CCCGGCT 1277
 OY 1553 GAGCAGCCCGGCGGCGGAG 1612
 DB 1278 GAGCCCTACCCCTCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1314
 OY 1613 CGCTCTCTGGCGGAG 1672
 DB 1315 -----ACCCCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1359
 OY 1673 GAGACCCGCGGAG 1732
 DB 1360 GGCAGCGCCCAAG 1413
 OY 1733 AGGAGCTGTCAACTCTCATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1792
 DB 1414 AGGAGCGGCTCACTCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1473
 OY 1793 AAGATGAGAGTCCCTACCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1852
 DB 1474 AAGCTGCAAGTTCGAG 1533
 OY 1853 CTGCAAAAG 1912
 DB 1534 CTGCGAAG 1593
 OY 1913 CTGCTGTAAAG 1972
 DB 1594 GTGGTCAATCAAG 1653
 OY 1973 TCGATCCCGAG 2032

DB 1654 TCGATTTCCCACTCTCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1713
 OY 2033 GCCAGTGGCGGCGGCGGCTGCTTCCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2092
 DB 1714 GCCAGGAGGAGGAG 1773
 OY 2093 TCTGAGGCTCAG 2152
 DB 1774 ACGGAGGCT-----GGGAGAGGCGCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1815
 OY 2153 GTGACCTTCACTCTCATCTGGGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2212
 DB 1816 GTCACTTCACTCTCATCTGGGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1875
 OY 2213 CAGCAGCAGCTCTGAGCAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2272
 DB 1876 CAGGCGCAGCTCTGAGCAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1935
 OY 2273 AACGG 2277
 DB 1936 AACGG 1940
 RESULT 6
 ABA08296
 ID ABA08296 standard; cDNA; 906 BP.
 XX
 AC ABA08296;
 XX
 DT 11-JUN-2002 (first entry)
 XX
 DE Human HRPOR-1 homologue-encoding cDNA, SEQ ID NO:72.
 XX
 KW Human: cytokine; cell proliferation; tissue growth; immunomodulator; growth factor;
 KW haematopoiesis; regulation; tissue growth; immunomodulator; growth factor;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; anti-inflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytoskeletal; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antilucer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157188-A2.
 XX
 PD 09-AUG-2001.
 XX
 PE 05-FEB-2001; 2001WO-0503800.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-45740/49.
 XX
 DR P-PSDB; ABB11052.
 XX
 PT Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 PS Claim 1; Page 364-365; 1963pp; English.
 CC Sequences ABB10981-ABB12310 represent 1350 novel human polypeptides, and

sequences ABA08225-ABA09574 represent nucleic acids encoding them. The invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides, antibodies against the polypeptides, methods of detecting the nucleotides or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; hematopoietic regulatory activity; tissue growth factor activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, hematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease or accidental damage. The polypeptides and nucleotides may also be used in the diagnosis of the above conditions, and in drug screening techniques. The present sequence represents a cDNA encoding a novel human polypeptide of the invention.

Sequence 906 BP; 205 A; 253 C; 267 G; 181 T; 0 other:

Query Match 22.1%; Score 658.8; DB 22; Length 906;
Best Local Similarity 87.9%; Pred. No. 1.3e-119;
Matches 786; Conservative 0; Mismatches 2; Indels 106; Gaps 2;

164 CAGGCCCAATATGTCGGCCCTATCGGCTGAGAGAGCGTCGGCAAGGACAGCAGG 223
1 CAGGCCCAATATGTCGGCCCTATCGGCTGAGAGAGCGTCGGCAAGGACAGCAGG 60
224 CTGTTAACTCGGGGTCACCTGCATCAGGAGTGAAGTCCGATCAATGCTGAAC 283
61 CTGTTAACTCGGGGTCACCTGCATCAGGAGTGAAGTCCGATCAATGCTGAAC 120
284 CCGGAGAAAGCTGTCGAGTGGTGTGATGAAGTGAAGCGGAGATCGCATCT-AAA 342
121 CCGGAGAAAGCTGTCGAGTGGTGTGATGAAGTGAAGCGGAGATCGCATCTCGAG 180
343 GCTCATCGAAGACACCATATCTCCAGAGTCCAGAGAGTCTACAGAGCAAGAAATATT 402
181 GCTCATCGAAGACACCATATCTCCAGAGTCCAGAGAGTCTACAGAGCAAGAAATATT 240
403 ----- 402
241 TCCCCCGGATGAATCAGATCAGTCCGTGATGCTGGCGAGTTTACACACAGGAAA 300
403 ----- GTACCTGGTCTGGA 417
301 GCTGTGCGCCGCGCTCATGGACCTGCTCTCAGGCTTTCACAGATCCTGTTCTGGA 360
418 GCACGCTCTGGGGGGGAGTATTCGACTACTCTGTAAGAAAGGAGAGACTGAGCCCAA 477
361 GCACGCTCTGGGGGGGAGTATTCGACTACTCTGTAAGAAAGGAGAGACTGAGCCCAA 420
478 GGAGGCCGAAATTTCTCGCGCAGATTTGTCGCGTGAAGTTCGCCAGAGTACTGC 537
|||||

421 GGAGGCCGAAAGTTCTTCCGCCAGATTTGTGTGCGCTGGACTTCCACAGCTACTC 480
538 CATCTCCACAGAGACTTAAGCCCGAAGACCTGCTTTGGATGAGAAAAACAATCCG 597
481 CATCTCCACAGAGACTTAAGCCCGAAGACCTGCTTTGGATGAGAAAAACAATCCG 540
598 CATTCGACATTTGGGATGCGCTCCCTGACAGTGGGGGACAGCTCTGGAGACACCTG 657
541 CATTCGACATTTGGGATGCGCTCCCTGACAGTGGGGGACAGCTCTGGAGACACCTG 600
658 CCGGTCCTCCCATTTATGCGGTCCAGAGTGAATTAAGGGGAAAAATATGATGGCCGG 717
601 CCGGTCCTCCCATTTATGCGGTCCAGAGTGAATTAAGGGGAAAAATATGATGGCCGG 660
718 GCGAGCATGTGGAGTGTGGAGTCACTCTTCGCCCTCTGTTGGGGCTGTGCCCTT 777
661 GCGAGCATGTGGAGTGTGGAGTCACTCTTCGCCCTCTGTTGGGGCTGTGCCCTT 720
778 TGATGACGACAACTCCGCGACGCTGCTGGAGAAAGTGAACGGGGCTCTTCCATGCG 837
721 TGATGACGACAACTCCGCGACGCTGCTGGAGAAAGTGAACGGGGCTCTTCCATGCG 780
838 CCACTTCATTCCTCCAGATGCGCAGAGCTCTGAGGGGAAATGCGAAGTGGAGCCGA 897
781 CCACTTCATTCCTCCAGATGCGCAGAGCTCTGAGGGGAAATGCGAAGTGGAGCCGA 840
898 AAAAGGCTCAGTCTGAGGCAAAATTCGAAACATCTTGTGACTAGCGGGAA 951
841 AAAAGGCTCAGTCTGAGGCAAAATTCGAAACATCTTGTGACTAGCGGGAA 894

RESULT 7
ABA08924
ID ABA08924 standard; cDNA: 614 BP.

ABA08924;

11-JAN-2002 (first entry)

Human Ser/Thr kinase homologue-encoding cDNA, SEQ ID NO:700.

Human: cytokine; cell proliferation; tissue growth; immunomodulator; growth factor;
haematopoietic regulation; tissue growth; immunomodulator; activin;
inhibin; chemotaxis; chemokinesis; thrombolytic; oncogenesis;
proliferation; metastasis; cancer; tumour; hematopoietic disorder;
myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
chronic inflammatory condition; proliferative retinopathy;
atherosclerosis; coronary heart disease; arterial ischaemia;
bone disorder; osteoporosis; vascular growth disorder;
tissue regeneration; wound healing; infection; immune disorder;
cell culture; drug screening; gene therapy; anti-inflammatory;
antiallergic; antiarthritic; haemostatic; antiarteriosclerotic;
cytostatic; osteopathic; vasotropic; cardiatic; virucide; antibacterial;
antifungal; vulnery; antilucer; ss.

Homo sapiens.

WO200157188-A2.

09-AUG-2001.

05-FEB-2001; 2001WO-US03800.

03-FEB-2000; 2000US-0496914.

27-APR-2000; 2000US-0560875.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT;

WPI; 2001-457740/49.

P-PSDB; AB11680.

proteins/T cell populations and antigen presenting cells expressing the polypeptide; methods for stimulating an immune response or treating cancer in a patient by administering an immune response or treating cancer comprising at least one of the oligonucleotide of, or an antibody and a detection reagent consisting of a reporter group. The polypeptides and polynucleotides are useful as vaccines for the treatment or prevention of lung cancer, and for diagnosis and monitoring of such cancer. The polynucleotide, polypeptide and antigen presenting cells can be used to stimulate or expand T cells specific for a tumorous protein. The polynucleotides may be used as probes or primers for nucleic acid hybridisation, and in the preparation of ribozyme molecules for inhibiting expression of tumour polypeptides and proteins in tumour cells. The present sequence is one of the 183 lung cancer associated polynucleotides.

50 Sequence 512 BP; 108 A; 149 C; 166 G; 89 T; 0 other;

Query Match	17.2%	Score 512;	DB 24;	Length 512;
Best Local Similarity	100.0%	Pred. No. 6.1e-91;		
Matches 512; Conservative	0;	Mismatches	0;	Gaps 0

QY	604	AGACTTCGCGATGGCTCCCTCCGTAGAGTGGGGGACACCTTCCTGGAGACACGCTGGGGGTC	663
Db	1	AGACTTCGCGATGGCTCCCTCCGTAGAGTGGGGGACACCTTCCTGGAGACACGCTGGGGGTC	60
QY	664	CCCCCATATGCTGTGCCAGAGGTATTAAGGGGGGAAAAATATGATGGCGCCGGGACAGA	723
Db	61	CCCCCATATGCTGTGCCAGAGGTATTAAGGGGGGAAAAATATGATGGCGCCGGGACAGA	120
QY	724	CATGTGAGCTGTGAGAGTATCCCTTCATGGCCGTCGAGGGGGGCTGCCCCTTGTGAGA	783
Db	121	CATGTGAGCTGTGAGAGTATCCCTTCATGGCCGTCGAGGGGGGCTGCCCCTTGTGAGA	180
QY	784	CGAACAACCTCCGCGACGCTGCTGGAGAAAGGTGAAGGGGGGCTTTCACATGCCCACTT	843
Db	181	CGAACAACCTCCGCGACGCTGCTGGAGAAAGGTGAAGGGGGGCTTTCACATGCCCACTT	240
QY	844	CATTCCGCCAGATTTGGCAGAGGCTCCTGAGGGGGAAATGATCGAAGTGGAGCCCGAAAAAG	903
Db	241	CATTCCGCCAGATTTGGCAGAGGCTCCTGAGGGGGAAATGATCGAAGTGGAGCCCGAAAAAG	300
QY	904	GCTCAGCTGTGAGCAAAATTCAGAAACATCTTGTAACCTAGGCGGGAAACAGAGCCAGA	963
Db	301	GCTCAGCTGTGAGCAAAATTCAGAAACATCTTGTAACCTAGGCGGGAAACAGAGCCAGA	360
QY	964	CCGTCGCTGTGAGCCAGCCCTCTGGCGCCGGGGTAAACCATGCGGAGCCCTGCATCCAAAG	1023
Db	361	CCGTCGCTGTGAGCCAGCCCTCTGGCGCCGGGGTAAACCATGCGGAGCCCTGCATCCAAAG	420
QY	1024	AGAGCTGTGAGCCCGAGAGTCTCTAGAGAGCATGTGACATCTGTGGGCTCTTCAGGGAGCCGGA	1083
Db	421	AGAGCTGTGAGCCCGAGAGTCTCTAGAGAGCATGTGACATCTGTGGGCTCTTCAGGGAGCCGGA	480
QY	1084	GAGGCTGCATCCCGAGAGTCTCGCAGTGTAGAGAG	1115
Db	481	GAGGCTGCATCCCGAGAGTCTCGCAGTGTAGAGAG	512

RESULT 9
ABL10489
ID ABL10489 standard; cDNA; 2720 BP.

DT	26-MAR-2002 (first entry)	
XX		
DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 25949	
XX		
XX	Drosophila: developmental biology; cell signalling; insecticide;	
KK	pharmaceutical; gene; ss.	
XX		
OS	Drosophila melanogaster.	

PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 XX
 PT 23-MAR-2001; 2001WO-US09231.
 XX
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX
 PA (PEKE) PE CORP NY.
 XX
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-6556860/75.
 DR P-PSDB; ABB66386.
 XX
 XX
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PT
 PS Claim 1; SEQ ID NO 25949; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from *Drosophila*. The invention is useful in developmental biology and in elucidating cell signaling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB1161176-AB130511), expressed DNA sequences (AB101840-AB1616175) and the encoded proteins (AB557737-AB572072).

The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

50 Sequence 2720 BP; 661 A; 715 C; 759 G; 585 T; 0 other;

Query Match	15.8%	Score 470.8;	DB 23;	Length 2720;
Best Local Similarity	66.4%	Pred. No. 1.1e-82;		
Matches 692;	Conservative 0;	Mismatches 347;	Indels 3;	Gaps 1

QY	162	AGCAGGCCCATATGATGCGCCCTCATGGCTGGAGAAAGACGCTGGCGAAGACAGACAG	222
Db	29	AGAAATGCGCAATTGTGTGGGGCCCATGCGCCCTGGAGAAAACCCCTTGGCAAGGCTCAAAACG	88
QY	222	GGCTGGTTAAACTGCGGGGTCCACTGCATCACGGGTCAAGAGGTCCGATCAAGATCGTGA	281
Db	89	GTCCTGCGCAAAATTTGGGGGTGCATTGTGTGATTTGGCAAGAAAGTTGCGATTAAATATCA	148
QY	282	ACCGGGAAGACGTCGTGCGAGTCGSGTGTGATGAAGAGTGAGACGGGAGATGCGCATCTCTGA	344
Db	149	ATCGCGAAGAAACTGACGGAAATCGGTGCTAATGAAAGTTTGAACTGGAATGCGCCATTAAATGA	208
QY	342	AGCTCATCGAAACCCACCATGTCTCAAGCTCCACGACGTCTACGAGAACAAAGAAATAT	401
Db	209	AACAAATCGATCATCCACACGCTCTTGCCCTGAGCCGATGTGTACGAGAACAAAGAAATAT	268
QY	402	TGTACGTGGTCTGAGACAGCTCTCGGGGGGTGAGCTATTCGACTACCTGGTAAAGAAAG	461
Db	269	TGTAATTTGATATTGGACATGATATCCGGCGAGAGCTCTTGATTAACCTGTGTGAAGAAAG	328
QY	462	GGAGACTGAGCGCCCAAGAGAGGCCGAAAGTCTTCCGCCAGATGTGTGCGCGTGACT	521
Db	329	GTCGATTGAGCGCGAAGAGAGGCGCGCAAGTTCTTCAGCGAAATCATCTCCGCCCTGGAAT	388
QY	522	TCTGCCACAGCTACTCATCTGCCACAGAACTAAAGCCGAGAACCTGCTTTTGGATG	581
Db	389	TCTGCCCACTCGCAATTCGATTTGGCCATGCGCACTTGAAGCGGAGAAATCTCTCGTGACG	448
QY	582	AGAAAAACAACATCCCGATGTCAGACTTCGCAATGGGTCCCGACAGGTGGGGAGACGCC	641
Db	449	AGGAAGATTAACATTAAAGATAGCGGACCTTTGGGATGCTTCCCTTCGACCCACGCTGGCGCA	508
QY	642	TCCGTGAGACCAAGCTGCGGGTCCCCCATTAATGCGTGTCCAGAGTCAATTAAAGGGGAAA	701

```

Db      509 TGTGGAGACCTCTCCGCGAGCCACATGAGGTGTCAGAGCTCATACGCGCGCAGA 568
Oy      702 AATATATGAGCCCGCCGAGCATGTGAGTGTGAGTATCTCTTCCGCTGCTCG 761
Db      569 AGTACATGCGCCGAGCGCGATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 628
Oy      762 TGGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 821
Db      629 TGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 688
Oy      822 GGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 881
Db      689 GGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 748
Oy      882 TCGAATGAGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 941
Db      749 TTAGAGTCAATCCGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 808
Oy      942 TAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1001
Db      809 CAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 865
Oy      1002 TGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1061
Db      866 CACACGTTATTCCTGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 925
Oy      1062 TGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1121
Db      926 TGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 985
Oy      1122 AAGAAAGATGATATATATATATATATATATATATATATATATATATAT 1181
Db      986 CGGAGAGGATATATATATATATATATATATATATATATATATATATATAT 1045
Oy      1182 ACCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1203
Db      1046 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1067

RESULT 10
AAK94923
ID      AAK94923 standard; cDNA; 2954 BP.
XX
AC      AAK94923;
XX
DT      06-NOV-2001 (first entry)
XX
DE      Human full-length cDNA, SEQ ID NO: 4157.
XX
KW      Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS      Homo sapiens.
XX
PN      EP1130094-A2.
XX
PD      05-SEP-2001.
XX
PE      07-JUL-2000; 2000EP-0114089.
XX
PR      08-JUL-1999; 99JP-0194486.
XX
PR      11-JAN-2000; 2000JP-0118774.
XX
PR      02-MAY-2000; 2000JP-0183765.
XX
PA      (HELI-) HELIX RES INST.
XX
PI      Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX      Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX      WPI: 2001-524255/58.
XX      P-PDB; AAM93956.
XX
PT      830 Primers useful for synthesizing full length cDNA clones and their

```

```

PR      use in genetic manipulation -
XX
PS      Claim 8; SEQ ID NO 4157; 1380bp + sequence listing; English.
XX
CC      The invention relates to primers for synthesizing full length cDNA
CC      clones. 830 cDNA molecules encoding a human protein have been
CC      isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC      molecules have been determined. Primers for synthesizing the full length
CC      cDNA are useful for clarifying the function of the protein encoded by
CC      the cDNA. The full length clones were obtained by construction of full
CC      length enriched cDNA libraries that were synthesised by the oligo-capping
CC      method. The primers enable the production of the full length cDNA easily
CC      without any special methods. The present sequence is a full length
CC      human cDNA of the invention.
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ      Sequence 2954 BP; 620 A; 970 C; 868 G; 496 T; 0 other;

Query Match          9.0%; Score 267.4; DB 22; Length 2954;
Best Local Similarity 55.6%; Pred. No. 7.6e-43;
Matches 514; Conservative 0; Mismatches 411; Indels 0; Gaps 0;

Oy      13 GTCCGCGGAGCGCTGGCGGAGCGCTGGCGGAGCGCTGGCGGAGCGCGTGGCGCG 72
Db      39 GTCTTCGCGGAGCGCTGGCGGAGCGCGGAGCGCGCTGGCGGAGCGCGTGGCGCG 98
Oy      73 GACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 132
Db      99 GGGCAGTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 158
Oy      133 CCTCCCGGAGCGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCG 192
Db      159 TGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 218
Oy      193 GGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 252
Db      219 GCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 278
Oy      253 GGGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 312
Db      279 TGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 338
Oy      313 GAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 372
Db      339 GAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 398
Oy      373 CCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 432
Db      399 GTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 458
Oy      433 TGAAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 492
Db      459 AGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 518
Oy      493 GTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 552
Db      519 GTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 578
Oy      553 CCTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 612
Db      579 CCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 638
Oy      613 CATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
Db      639 CTTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 698
Oy      673 TGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 732
Db      699 TGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 758
Oy      733 CTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 792

```


DT 27-MAR-2001 (first entry)
 XX Novel protein kinase cDNA, SEQ ID NO: 35.
 DE
 XX Human: mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
 KW immunosuppressive; cardiantic; renal; antiinflammatory; antiasthmatic;
 KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;
 KW immune disorder; cardiovascular disease; neurodegenerative disease;
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.
 XX
 XX Mus musculus.
 XX
 XX W0200073469-A2.
 XX
 XX PN 07-DEC-2000.
 XX
 XX PD 26-MAY-2000; 2000MO-US14842.
 XX
 XX PF 28-MAY-1999; 99US-0136503.
 XX
 XX PR (SUGEN-) SUGEN INC.
 XX
 XX PA Plowman GD, Martinez R, Whyte D, Sudersanam S;
 XX PI P-PSDB; AAB55629.
 XX
 XX DR Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 XX PT treating immune-related diseases and disorders, cardiovascular disease,
 XX PT neurodegenerative diseases and/or cancers -
 XX
 XX PS Disclosure; Fig 2; 310pp; English.
 XX
 XX CC The present sequence encodes a novel protein kinase. The nucleic acids
 XX CC and the protein kinases they encode may be used in the treatment and
 XX CC diagnosis of diseases associated with inappropriate kinase expression
 XX CC such as immune-related diseases and disorders, cardiovascular disease,
 XX CC neurodegenerative diseases and/or cancers. The nucleic acids and
 XX CC complementary sequences may also be used as DNA probes in diagnostic
 XX CC assays. The kinase polypeptides may be used as antigens in the production
 XX CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 XX CC and kinase antagonists may also be used to down regulate kinase
 XX CC expression and activity. Diseases related to kinase expression and
 XX CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 XX CC disorders, complications of organ transplantation, myocardial infarction,
 XX CC immune disorders, cardiomyopathies, strokes, renal failure,
 XX CC oxidative-stress related disorders, chronic inflammatory bowel disease,
 XX CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
 XX CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
 XX CC reproductive disorders.
 XX
 XX SQ Sequence 1594 BP; 376 A; 488 C; 458 G; 272 T; 0 other;
 XX
 XX Query Match 8.88; Score 261.6; DB 22; Length 1594;
 XX Best Local Similarity 58.98; Pred. No. 8.9e-42;
 XX Matches 450; Conservative 0; Mismatches 314; Indels 0; Gaps 0;
 XX
 XX QY 174 ATGTGGCCCTATCGCTGAGAGAGAGCTGGGCAAGAGACAGGCTGTTAAAC 233
 XX DB 80 ATGTGGCAATATAGGCTGCTAAGGACCATCGGAGGCAACTTCGCCAAGTCAAC 139
 XX QY 234 TCGGGCTCCTGCTGATCAGAGGCTCAGAGGCTCCATCAGATCGTGAACCGGAGAAC 293
 XX DB 140 TGGCTCGGCAATATCTCTCGGCGGCGGAGTCTTAAGATTAAGTAAAGCCAC 199
 XX QY 294 TGTGGAGTGTCTGCTGATGAAGTGAAGGAGCGGAGATGCCCTCGAAGCTCAGCAAC 353
 XX DB 200 TGAACCCAGTAGCTTGAGAGAGCTGTCAGAGAACTCCGATTTAGAGGAGCAATACC 259
 XX QY 354 ACCCATGTCCTCAGAGCTCCAGAGCTCTACAGAGAAAGAAATATTTGATCTGCTGC 413
 XX DB 260 ACCCAACATGCTGAAGCTTTTGAAGTATAGAGACGAGAGAAAGCCTATACCTGCTGA 319

QY 414 TGGACAGCTCTCGGGGGGTGAGCTATTCAGACTACCTGGTAAAGAGGGGAGACTGACGC 473
 DB 320 TGGATATACGCTAGGCGCAGAGAGAGTGTTCAGTACTCTGTGCGACGCGCCATGAAAG 379
 QY 474 CCAAGAGAGCCCGAAGTCTTCGCGCAGATTTGCTGCGCTGAGACTTCCACAGCT 533
 DB 380 AGAAGAGAGCGCTCGAGCCAGATTCGCGGAGAGTCGTACGCCGTGACACTACTATCAGA 439
 QY 534 ACTCATCTGCGCAGAGACTTAAAGCCGAGAACCTGCTTTGGATGAGAAAAACA 593
 DB 440 AGAATATTGTACACAGGAGATCTAAAGGCTGAAGAACTGTCTGCTGATCCGAGCCACA 499
 QY 594 TCCGATTTGAGACTTGGGCTGGGCTGCCGAGCTGGGGAGACGCTCTCGAGACA 653
 DB 500 TCAAAATGCGGACTTGGCTGCGCTCAGCAATGAGTTACGCTGGGCTCAAGCTGGACACT 559
 QY 654 GCTGCGGGTCCCGCCATTTATCGTGTCCAGAGTGATTAAGGGGAAAAATATGATGGC 713
 DB 560 TCTGTGGAGAGCCCGCCATACGCGCCCGCAGAGCTGTCCAGGCGAAGATGATGGC 619
 QY 714 GCGGGCAGACATGTGAGCTGTGAGTCAATCTCTGCGCCCTGCTGTGGGGCTCTGC 773
 DB 620 CAGAGGTGACATGTGAGCTGTGAGTCAATCTCTGAGAGAGAGAGAGAGAGAGAGAG 679
 QY 774 CTTTGTATGAGAGACACTCCGCGCAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 833
 DB 680 CTTTGATGGGACACACTCAAGAGCTGCGGAGCGAGTCTCAGAGAGAGAGAGAGAGAG 739
 QY 834 TGGCCCTTCTATCTCCAGATTCGCGAGAGCTGCTGAGGAGATGATGCAAGTGAAGC 893
 DB 740 TCCCTTCTACATCTCTACAGACTGCGAGAGACTTCTCGGAGATTTCTGCTGTAAC 799
 QY 894 CCGAAAAAGGCTGCTGAGGCAAAATTCAGAAATCTCTTG 937
 DB 800 CCGCAAAAGGCTGACCTGAGAGCAAAATTCAGAAATATG 843
 RESULT 15
 ID ABA05737 standard; cDNA: 3170 BP.
 AC ABA05737;
 XX
 XX DT 04-MAR-2002 (first entry)
 XX
 XX DE Murine neuronal serine threonine protein kinase 9B5 coding sequence.
 XX
 XX KW Mouse; neuronal serine threonine protein kinase; apoplexy; cytostatic;
 KW cerebrioprotective; neuroprotective; antiparkinsonian; immunosuppressive;
 KW virucide; anti-HIV; antiarthritic; antiinflammatory; antileukemic; cancer;
 KW cerebral infarction; neurological disease; immunological disease; atopy;
 KW infection; leukaemia; polyarthritits; Crohn's disease; ulcerative colitis;
 KW 9B5; ss.
 XX
 XX OS Mus musculus.
 XX
 XX XX Key Location/Qualifiers
 XX FH 1..2175
 XX FT CDS /-tag- a
 XX FT /product- "9B5"
 XX FT 3094..3099
 XX FT polyA_signal
 XX FT /-tag- b
 XX
 XX PD W020018108-A1.
 XX
 XX DB 22-NOV-2001.
 XX
 XX PF 17-MAY-2001; 2001MO-EP05660.
 XX
 XX PR 17-MAY-2000; 2000DB-1024171.
 XX
 XX PA (BADI) BASF-LYNN BIOSCIENCE AG.

XX Schneider A, Klausner B, Fischer A, Newrzella D, Goetz B;
 PI Rosser M, Eisenhardt G, Kuner R, Trutzel A, Kammandel B;
 PI Jomana Naim S, Schwaninger M;
 XX
 DR MPI: 2002-055696/07.
 P-PSDB: ABB04431.

XX New polynucleotide, useful for the diagnosis, treatment and prevention
 PT of e.g. apoptosis, tumors and autoimmune disease, comprises a
 PR polynucleotide encoding the neuronal protein kinase -

PS Claim 1: Page 52-53; 75pp: German.

CC The present invention provides the protein and coding sequences of two
 CC neuronal serine threonine protein kinases each from the mouse and human.
 CC These are designated 985 and 985.B. Detection of their expression levels
 CC can be used to diagnose the risk of apoptosis and in prognosis of cerebral
 CC infarction. They can also be used to identify agents useful in the
 CC prevention, treatment and diagnosis of neurological diseases, including
 CC stroke, multiple sclerosis, Parkinson's disease, amyotrophic lateral
 CC sclerosis, tumours, especially carcinoma, immunological disorders,
 CC including autoimmune diseases, atopy, viral (including human immune
 CC deficiency virus) infections, acute or chronic leukocytic or myeloid
 CC leukaemia, primary chronic polyarthritis, Crohn's disease and ulcerative
 CC colitis. The present sequence is the murine 985 coding sequence.

XX Sequence 3170 BP; 714 A; 988 C; 897 G; 571 T; 0 other;

Query Match 8.7%; Score 259.4; DB 24; Length 3170;

Best Local Similarity 57.4%; Pred. No. 2.9e-41; Matches 467; Conservative 0; Mismatches 346; Indels 0; Gaps 0;

QY 174 ATGTGGGCCCCCTATCGCTGAGAGACGCTGGCCAAAGACAGAGCGCTGTAAAC 233
 DB 80 ATGTGGGCACTATAGCTGTGTAGACCATCGGGAAGGCACTTGCCTCAAGTCAAGC 139
 QY 234 TCGGGGTCCACTGTCATCAGGGGTGAGAGGTCCCATCAAGTCTGAAACGGGAGAAC 293
 DB 140 TGGCTGGGCTATTCCTCAGCGGCGGAGGCTGCTATTAAGATCATTAAGACCCAGC 199
 QY 294 TGTGGAGTGTGCTGATGAGAGGTGAGGAGAGATGCGGCTGCTGCTGCTGCTGCTG 353
 DB 200 TGAACCCCACTAGCTTCAAGAGCTGTTCAGAGAGATCCGATTAAGAGGACTCAAC 259
 QY 354 ACCCATGTCCTCAAGCTCCAGCAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 413
 DB 260 ACCCAACATCGTGAAGCTTTTGAAGTGAATAGAGAGAGAGAGAGAGAGAGAGAGAG 319
 QY 414 TGAAGACGCTGCGGGGGGTGAGTATTCAGTACCTGTGTAAGAGAGGAGACTGACGC 473
 DB 320 TGAATACGCTAGCGGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 379
 QY 474 CCAAGAGAGGCGGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 533
 DB 380 AGAAGAGAGGCTCAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 439
 QY 534 ACTCATCTGCAAG 593
 DB 440 AGAACAATTGACAG 499
 QY 594 TCGCATTTGAGAGCTTGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 653
 DB 500 TCAAAATCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 559
 QY 654 GGTGGGGTCCCGCATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 713
 DB 560 TCTGTGGAG 619
 QY 714 GCGGAG 773
 DB 620 CAGAGGTGAG 679

QY 774 CTTTGTAGACGACACACCTCCCGAGCTGCTGAGAGAGTGAAGAGAGAGAGAGAGAG 833
 DB 680 CTTTCATGGGACACACCTCAAGAGAGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 739
 QY 834 TGGCCCATCTCATCTCCAGATTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 893
 DB 740 TCCCTTCTACATGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 799
 QY 894 CCGAAAAAGGCTCAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 953
 DB 800 CCGCAAAAGCTGTACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 859
 QY 954 ACGAGCCAG 986
 DB 860 AGGAG 892

Search completed: April 21, 2003, 20:31:16
 Job time : 428.371 secs

Db	269	ATTTTGACCGTGGTCTAAGAACAGCTGTACAGTGTGTAGCTCTTTGACACTACTCGTGGAAGA	328
OY	459	AGGGAGACTGACGCCAMAGAGGCCCGGAAAGTTCTTCCGCCAGATGTGTGTGGCTGG	518
Db	329	AGGGAGGCTGACGCCCTAAGAGAGGCTGGAAAGTTCTTCCGCCAATCATCTCTGGGCTGG	388
OY	519	ACTTTCGCCACAGCTACTCTCATCTGTCCACAGAGACCTAAAGCCCGAGAACCTGCTTTTGG	578
Db	389	ACTTTCGCCACAGCCACTCTCATATGCCCACAGGGATCTGAAACCTGTGAACCTCTGCTGG	448
OY	579	ATGACAAAACAAACATCCGACTTGCAGACACTTGGCATGGCGTCCCTCGAGGTGGGGACA	638
Db	449	ACGAGAGAAACAAACATCCGATCGCATGGCACTTTGGCATAGGCGTCCCTCGAGTTGGCGACA	508
OY	639	GCCTCTGGAGACCACTGCGGGTCCGCCCATTAATGCGTGTCCAGAGGTGATTAAAGGGGG	698
Db	509	GCCTGTGGAGACCACTGTGGGTCCGCCCATTAATGCGTGTCCAGAGGTGATTAAAGGGGG	568
OY	699	AAAAATATATATGCGCCCGCGGGCAGACATGTGGAGCTGTGGAGTATCTCTTCCGCCCTGC	758
Db	569	AGAAATATATACGGCCCGGAAAGCGGAGCTGTGGAGCTCGGCGTCTATCTGTTCGCCTTTCG	628
OY	759	TGCTGGGGGCTCTGCGCTTGTATGACAGACAACTCCGCCAGCTCTGTGAGAAAGGTGAAC	818
Db	629	TGCTGGGGGCTCTGCGCTTGTATGACAGACAACTTTCGCACAGCTCTGTGAGAAAGGTGAAC	688
OY	819	GGGGGCTCTTCCACATGCCCATGTCCTTATTCCTCCAGATTGCCAGAGGCTCTCTGAAGGGAA	878
Db	689	GGGGGCTGTTCACATGCGCGCATTTATCCCGCCGACTGCCAGAGTGTCTACGGGGCA	748
OY	879	TGATGSAAGTGGAGCCCGGAAAAAGCTCAGTGTGGAGCAAAATTCAAAACATCTTGTGT	938
Db	749	TGATGSAAGTGGAGCGCCGACGCGCGCTCAGGTAGAGCAATTCAAAAACACATATGTGT	808
OY	939	ACCTAGGCGGGAAACAGAGCCAGACCCGTGCTGGAGCCAGCCCTGGCGCGCGGGGTAG	998
Db	809	AATATAGGGGGCAAGATATGAGCCCGGAAC-----AGAGCAGCCCATTTCTGTGCAAGGTGC	862
OY	999	CCATGCGGAGCCTGCCATCCACAGAGAGCTGGAGCCCGGACGTCCTAGAGACATGGCAT	1058
Db	863	AGATCCGCTGCGTGGCCAGCCTTGAGAGACATGCAACCCCGAGTGTGAGACATGTCAC	922
OY	1059	CACGTGGGCTGTTCACAGGAGCCGGAGAGGCTGCATCGCGAGCTGCGCAGTGTAGAGAGAGA	1118
Db	923	CACGTGGGCTGTTCACAGGAGCCGCAACAGGTGTGTGCAAGACCTCTGTCCGAGAGAGAGA	982
OY	1119	ACCAAGAAAGATGATATATATCTGTTTGGATGGAAGAGCGGTATGCCAGCTGTGT	1178
Db	983	ACCAGAGAGAGATGATTACTTCTCTCTCTGACCGGAAAGAAAGGTATCCCGAGCCAGG	1042
OY	1179	AGGACCGAGACCTGCTCCCGGAAATGATGTTGACCCCGCCCGGAACGTTGTGATTCTC	1238
Db	1043	AGGATGAGAGACCTGCCCGCCCGGAAAGAGATATACCTCTCCCGGAACGTTGTGACTCTCC	1102
OY	1239	CCATGCTGAGCCGTACAGGAGAGCGCGGACACAGAGGAAATGCTATGAAATCTTGACA	1298
Db	1103	CGATCTGTAAACCGGACCGGCAACGCGCGGCGCAACAGCAATCATGAGGTGTCTCAGCG	1162
OY	1299	TCACCGATGTCGGGGGTGGTGGCTCCCTGTACCAACCGGAGCGGCTTGGAGATGGGCC	1358
Db	1163	TGAC-----GGACGGCGGCTCCCGGCTGCTGGCGCGCGCGGCAATGTAGATGGGCC	1213
OY	1359	AGCAGCGCAGATATCCGCTAGCGTCAAGTGAAGCTTCAAGGGGTCTGCTCCAGCCTC	1418
Db	1214	AGCAGCGCAGAGGTCTCGGTTCATAGCGGTGTCTCTCAAGGCTTTTCCACCGACCCAC	1273
OY	1419	TAAAGACCCCAAGAGTCCGGTCTTTTCTTTTTCACCGAGCCGGGGGCTGTGAGATGAGG	1478
Db	1274	TCAGCAGCCCGC-----	1286
OY	1479	CTCGAGGCGGGGGCTCCCGGACTTCCAAAACGAGAGCGTGTCTTCTTCGGGGCCCGCAGG	1538

[illegible]

```

? RESULT 2
? US-09-842-582-3
? Sequence 3, Application US/09842582
? Patent No. US20020155570A1
? GENERAL INFORMATION:
? APPLICANT: Millennium Pharmaceuticals, Inc.
? APPLICANT: Meyers, Rachel
? TITLE OF INVENTION: 2246, NOVEL PROTEIN KINASE MOLECULES AND
? TITLE OF INVENTION: US6, THEREFOR
? FILE REFERENCE: 38155-20054.00
? CURRENT APPLICATION NUMBER: US/09/842,582
? CURRENT FILING DATE: 2001-04-25
? PRIOR APPLICATION NUMBER: US 60/159,391
? PRIOR FILING DATE: 2000-04-25
? NUMBER OF SEQ ID NOS: 11
? SOFTWARE: FASTSD for Windows Version 4.0
? SEQ ID NO 3
? LENGTH: 2025

```

TYPE: DNA
ORGANISM: Homo sapiens
US-09-842-582-3

Query Match 32.2%; Score 961; DB 9; Length 2025;
Best Local Similarity 70.1%; Pred. No. 1,8e-215;
Matches 1477; Conservative 0; Mismatches 435; Indels 195; Gaps 6;

159 CCCGACGACGCAATATGTTGGGCCCCCTATGCGGTGAGAAAGAGCTGGGCAAGAGCAGA 218
DB CCGAGCAGCGCGAGTATGTTGGGCCCCCTATGCGGTGAGAAAGAGCTGGGCAAGAGCAGA 88
219 CAGGGGTGTTAACTCCGCGGTCCATGCAAGGCTCAAGAGTCCCTCAAGATCG 278
DB CAGGTGTGGTAACTCCGCGGTCCATGCAAGGCTCAAGAGTCCCTCAAGATCG 148
279 TGAACCGGAGAACTGTGCGAGTGGTGTGATGAGTGGAGCGGAGATCGGCATCC 338
DB TCAACCGTGAAGAGCTCAGGAGTGGTGTGATGAGTGGAGCGGAGATCGGCATCC 208
339 TGAAGCTCATGCAACACCAATGCTCTCAAGCTCAAGAGTCTACGAGAACAAAGAT 398
DB TGAAGCTCATGCAACACCAATGCTCTCAAGCTCAAGAGTCTACGAGAACAAAGAT 268
399 ATTGTGACCTGTTCTGAGACAGCTCTCGGGGGGTGAGCTATGAGTACTGCTGTAAGA 458
DB ATTGTGACCTGTTCTGAGACAGCTCTCGGGGGGTGAGCTATGAGTACTGCTGTAAGA 328
459 AGGAGGAGACTGAGCGCCAGAGGAGCGCGAAAGTCTTCCGCGAGATTGCTGCGCTGG 518
DB AGGAGGAGCTGAGCGCCAGAGGAGCGCGAAAGTCTTCCGCGAGATTGCTGCGCTGG 388
519 ACTTGTCCACAGCTACTCATCTGTCAGAGAGACTTAAAGCCCGAGAGACTGCTTTGG 578
DB ACTTGTCCACAGCTACTCATCTGTCAGAGAGACTTAAAGCCCGAGAGACTGCTTTGG 448
579 ATGAGAAAAAACAATCCGATGCAAGATCGGATGCGCTCCCTGAGAGTGGGGAGCA 638
DB ATGAGAAAAAACAATCCGATGCAAGATCGGATGCGCTCCCTGAGAGTGGGGAGCA 508
639 GCGTCTGAGACAGCTGCGGGTCCCCCAATATGCGTGTCAAGAGTGAATTAAGGGGG 698
DB GCGTGTGAGACAGCTGCGGGTCCCCCAATATGCGTGTCAAGAGTGAATTAAGGGGG 568
699 AAAAATATGATGAGCCCGCGGAGACATGTGAGCTGTGAGTCAATCTTCCGCTTGC 758
DB AAAAATATGAGCCCGCGGAGAGCGGAGCTGTGAGCTGTGAGTCAATCTTCCGCTTGC 628
759 TCGTGGGGGCTGCGCTTGTGATGACAAACCTCCGCGAGCTGTGAGAGAGTGAAC 818
DB TCGTGGGGGCTGCGCTTGTGATGACAAACCTCCGCGAGCTGTGAGAGAGTGAAC 688
819 GGGGCTCTTCAACATGCGCCACTTCACTCCCTCAGATTGCGAGAGCTCTGAGAGGGA 878
DB GGGGCTGTCTCAACATGCGCCACTTCACTCCCTCAGATTGCGAGAGCTCTGAGAGGGA 748
879 TGAATGATGAGAGCCCGGAAAGAGCTCAATGCTGAGCAAAATTCAGAAACATCTTGGT 938
DB TGAATGATGAGAGCCCGGAAAGAGCTCAATGCTGAGCAAAATTCAGAAACATCTTGGT 808
939 ACTTAGCGGGGAAACAGAGACCAACCGGCTGAGAGCCGCGCTGCGCGGGGATG 998
DB ACTTAGCGGGGAAACAGAGACCAACCGGCTGAGAGCCGCGCTGCGCGGGGATG 862
999 CCATGCGGAGCTGCTCAATCAACAGAGAGCTGAGACCGCGAGCTCTAGAGAGCATGGCAT 1058
DB AGATTCGCGTGTGCTGCGAGCTGAGAGAGCATGAGACCGCGAGCTGTGAGAGAGCATG 922
1059 CACTGGGCTCTTCAAGAGACCGGAGAGGCTGATCGGAGCTGCGGAGTGAAGAGAGA 1118
DB CACTGGGCTCTTCAAGAGACCGGAGAGGCTGATCGGAGCTGCGGAGTGAAGAGAGA 982
1119 ACAAGAAAGATGATATATATCTGCTTTGATCGGAAGAGCGGATATCCAGCTGTG 1178

983 ACCAGAGAGATGATATATATCTGCTTTGATCGGAAGAGCGGATATCCAGCTGTG 1042
1179 AGGAGCAGACCTGCTCCCGGAAATGATGTTGACCCCCCGGAGAGCTGTGATCTC 1238
DB AGATATGAGAGCTGCCCCCGGAGAGCATGATGACCTCCCGGAGAGCTGTGATCTC 1102
1239 CCATGCTGAGCCGTCAACGGAGCGCGGAGAGCGGAGGAAATTCATGAAAGTCTGAGCA 1298
DB CGATGCTGAAACCGGAGAGCGGAGAGCGGAGGAAATTCATGAAAGTCTGAGCG 1162
1299 TCAACCATGCGGGGGGTGGTCCCTGTAACCCAGCCGAGCGGCTTGAAGATGCGCC 1358
DB TCAAC-----GAGGGGGGTGGTCCCTGTAACCCAGCCGAGCGGCTTGAAGATGCGCC 1213
1359 AGCAGACGAGATCCCTGAGTCAAGTGAAGCTCTCAGCGGTCTGCTCCAGAGCTC 1418
DB AGCAGCGCAGAGTCTCGGTCCATGAGCGGTGCTCTCAGAGCTTCCAGAGCGCAC 1273
1419 TAAAGACCCCAAGAGTCCGCTCTTTTCAACCGGAGCCGGGGCTGAGATGAGG 1478
DB TAAAGACCCCGC----- 1286
1479 CTGAGAGCGGGGCTCCCGACCTTCAAAACGAGAGCTGCTTCTCGGGGCCCGAGG 1538
DB CTGAGAGCGGGGCTCCCGACCTTCAAAACGAGAGCTGCTTCTCGGGGCCCGAGG 1286
1539 GTGGGGGCGCGGGAGACACCCCGCCCAAGTGCCTGCACACCCCTGCGCGGC 1598
DB GTGGGGGCGCGGGAGACACCCCGCCCAAGTGCCTGCACACCCCTGCGCGGC 1326
1599 CCGGAGCTCCCGCGCTCTGCGGAGACCCCTTGCATGCTGTCAGACAGCCGC 1658
DB CCGGAGCTCCCGCGCTCTGCGGAGACCCCTTGCATGCTGTCAGACAGCCGC 1357
1659 GGGCAGCTCCAGCGGAGACCCCGGGAGCAACACCCCGCGCGCGGTGGCGTCG 1718
DB GGGCAGCTCCAGCGGAGACCCCGGGAGCAACACCCCGCGCGCGGTGGCGTCG 1411
1719 GGGGAGCCGCTGAGAGAGTGTCTCAATCCATCCGAGAGCTCTGCGGCTCCCTC 1778
DB GGGGAGCCGCTGAGAGAGTGTCTCAATCCATCCGAGAGCTCTGCGGCTCCCTC 1471
1779 GCTTCAACGGGCGCAAGAGTCCCTCAACCTGAGAGATGTCAGCTTGAAGCCAG 1838
DB GCTTCAACGGGCGCAAGAGTCCCTCAACCTGAGAGATGTCAGCTTGAAGCCAG 1531
1839 AGTCCATCCCGGAGCTGGGAAACGCTCGGTCCGGGAATTCATCTCTTGGAGCAAG 1898
DB AGTCCATCCCGGAGCTGGGAAACGCTCGGTCCGGGAATTCATCTCTTGGAGCAAG 1591
1899 AAGAAATATATCTGCTCTTAAAGAGCAAACTCTCAGAGAGCTCAAAAGAGATCG 1958
DB AAGAGAGATCTCTGCTGATCAAAAGAGCAAACTCTGAGTCTCAAGAGTCAAGATCG 1651
1959 TCCATGCTTCTGCTGATCCCAAGCTGATGATGATGATGATGATGATGATGATGATGAT 2018
DB TCCATGCTTCTGCTGATCCCAAGCTGATGATGATGATGATGATGATGATGATGATGAT 1711
2019 GGGCGAGTCAAGGAGGAGGCGCGCTCCGCTTCCAAACCGCTGCGCTTCCAGG 2078
DB GGGCGAGTCAAGGAGGAGGCGCGCTCCGCTTCCAAACCGCTGCGCTTCCAGG 1771
2079 TGAACATCAAGCTCTCTGAGAGGTCAGAGCCCTCCCGGAGAGGAGCGGAGCGGAGG 2138
DB TGAATATCAACTACAGAGAGG-----GGGAGAGCGGAGAGAGAGGA 1813
2139 GTGCAATCTACTGCTCACTTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 2198
DB ACAGCAATCTACTGCTCACTTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 1873
2199 TGGTGAAGACATCAAGGAGAGTCTGAGACATCAAGACCAAGCCCTCGTGAAGGCC 2258
DB TGGTGAAGACATCAAGGAGAGTCTGAGACATCAAGACCAAGCCCTCGTGAAGGCC 1111

Db 1874 TGTGGAGACATCCAGGCCGCTGCTGAGCAGACAGACCCGCTGCGGCCAGCACT 1933
QY 2259 TGGCAGA 2265
Db 1934 TGTGAGA 1940

RESULT 3

US-09-842-582-1
Sequence 1, Application US/09842582
Patent No. US2002015570A1

GENERAL INFORMATION:

APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Meyers, Rachel
TITLE OF INVENTION: 2246, NOVEL PROTEIN KINASE MOLECULES AND
FILE OF INVENTION: USES THEREFOR
FILE REFERENCE: 38155-20054.00
CURRENT APPLICATION NUMBER: US/09/842,582
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: US 60/199,391
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2219
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (53)...(2077)
US-09-842-582-1

Query Match 32.2%; Score 961; DB 9; Length 2219;
Best Local Similarity 70.1%; Pred. No. 1.9e-215;
Matches 1477; Conservative 0; Mismatches 435; Indels 195; Gaps 6;

QY 139 CCCACAGCCCAATATGCGGCCCCATATCGGTGAGAGAGAGCGTGGCAAGAGAGA 218
Db 81 CGACACAGCCGAGATATGTTGGCCCTACCGGTGAGAGAGAGCGTGGCAAGAGAGA 140
QY 219 CAGGCGTGTAACTGCGGCTGCTGATGATGATGATGATGATGATGATGATGATG 278
Db 141 CAGGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 200
QY 279 TGAACCGGAGAGCTGTGAGAGCGTGTGATGATGATGATGATGATGATGATGATG 338
Db 201 TCAACCGGAGAGCTGTGAGAGCGTGTGATGATGATGATGATGATGATGATGATG 260
QY 339 TGAACCTATGACACACCCATGCTCAAGCTCCAGCGCTCTAGAGAGAGAGAGAT 398
Db 261 TGAACCTATGACACACCCATGCTCAAGCTCCAGCGCTCTAGAGAGAGAGAGAT 320
QY 399 ATTGTACCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 458
Db 321 ATTGTACCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 380
QY 459 AGGGAG 518
Db 381 AGGGAG 440
QY 519 ACTTGTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 578
Db 441 ACTTGTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 500
QY 579 ATGAG 638
Db 501 ATGAG 560
QY 639 GCGTCTGAG 698
Db 561 GCGTCTGAG 620
QY 699 AAAAATATGATGCGCGCGCGCGCGAGAGATGATGATGATGATGATGATGATGATG 758

Db 621 AGAAGTATGACCGCGGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 680
QY 759 TCGTGGGGGCTGCGCTTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 818
Db 681 TGTGGGGGCTGCGCTTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 740
QY 819 GGGGCTGCTTCCACATGCGCCCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 878
Db 741 GGGGCTGCTTCCACATGCGCCCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 800
QY 879 TGTATGAG 938
Db 801 TGTATGAG 860
QY 939 ACCTAG 998
Db 861 ATATAG 914
QY 999 CCATGCGGAG 1058
Db 915 AGATCGGCTGCTGCGCTTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 974
QY 1059 CACTGCGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1118
Db 975 CACTGCGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1034
QY 1119 ACCAAG 1178
Db 1035 ACCAAG 1094
QY 1179 AG 1238
Db 1095 AG 1154
QY 1239 CCATGCGGAG 1298
Db 1155 CCATGCGGAG 1214
QY 1299 TCACGAG 1358
Db 1215 TCACGAG 1265
QY 1359 AGCAG 1418
Db 1265 AGCAG 1325
QY 1419 TAAAG 1478
Db 1326 TAAAG 1338
QY 1479 CTCGAG 1538
Db 1339 CTCGAG 1338
QY 1539 GTGGGGGCGCGGAG 1598
Db 1339 GTGGGGGCGCGGAG 1378
QY 1599 CCGCAG 1658
Db 1379 CCGCAG 1409
QY 1659 GGGCAGTCCACCGGAG 1718
Db 1410 GGGCAGTCCACCGGAG 1463
QY 1719 GGGGAGCGGCTGAG 1778
Db 1464 GGGGAGCGGCTGAG 1523
QY 1779 GCTTACACCGGAG 1838

```

Db 1524 GCTTCACCCGCGAAGCTGCAAGTTCGACGCGGAGAGATGTCACACCTGACACAG 1583
QY 1839 AGTCTCCCGGAGAGTGGCAAAAGCTCTGTGGGAGACTTCATCTCTTGAGCAAG 1598
Db 1584 AGTGTGCTCCCGAGAGTGGCAAAAGTCTGTGGGAGACTTCATCTCTTGAGCAAG 1643
QY 1899 AAGAACAAATATTCCTGTGGTAAAGCAAAACCTGTGAGCATCAAGCAAGAGATG 1558
Db 1644 AGGAGCAGATTTCTGTGTCATCAAGCAAAACCTGTGAGCATCAAGAGGTCATG 1703
QY 1959 TCATGCTTCTGTGATCCCAAGCTGTAGTCAAGTGTGTGTCACAGACACTTCA 2018
Db 1704 TGACGCTCTCTGTGATCCCAAGCTGTAGTCAAGCTGTGTGTCACAGACACTTCA 1763
QY 2019 GGGCGGAGTCAAGAGCCAGTGGGCGCCCTCGTCTTCCAAAAGCCGTCCTCCAG 2078
Db 1764 GGGCGGAGTCAAGAGCCAGTGGGCGCCCTCGTCTTCCAAAAGCCGTCCTCCAG 1823
QY 2079 TGACATGAGCTCTGTGAGGATCCAGAGCCCTCCCGGAGGAGGAGGAGGAG 2138
Db 1824 TTGATATCACTTACAGGAGGT-----GGGAGGCGCAGAGAGAGA 1865
QY 2139 GTGACATCACTCTCGTCACTTCACTCTCACTCTCGGAGTCCAGCCGTCGATCAAG 2198
Db 1866 AGGAGCATCTCTCGTCACTCTCACTCTCACTCTCGGAGTCCAGCCGTCGATCAAG 1925
QY 2199 TGTGTGAGACATCCAGGACAGCTCTGTGAGCATCTCAAGCCCTCCGTCAGAGCC 2258
Db 1926 TGTGTGAGACATCCAGGACAGCTCTGTGAGCATCTCAAGCCCTCCGTCAGAGCT 1985
QY 2259 TGGCAGA 2265
Db 1986 TGTGAGA 1992

```

RESULT 4
US-10-054-579-3
; Sequence 3, Application US/10054579
; Patent No. US20020137913A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; TITLE OF INVENTION: No. US20020137913A1 Human Kinases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0300-USA
; CURRENT APPLICATION NUMBER: US/10/054, 579
; PRIOR APPLICATION NUMBER: 2002-01-22
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-054-579-3

Query Match 28.6%; Score 851.4; DB 12; Length 1827;
Best Local Similarity 68.9%; Pred. No. 8.5e-190;
Matches 1356; Conservative 0; Mismatches 416; Indels 195; Gaps 6;

```

QY 311 ATGAAGTGGAGCGGAGATGCGCATCTCTGAAGCTCATGGAACCCCAATGCTCAAG 370
Db 1 ATGAAGTGGAGCGGAGATGCGCATCTCTGAAGCTCATGGAACCCCAATGCTCAAG 60
QY 371 CTCACGAGCTCTAGAGAACAAATATTTGTACTGTTCTTGAGACAGCTGTGGGG 430
Db 61 CTCACGAGCTCTAGAGAACAAATATTTGTACTGTTCTTGAGACAGCTGTGGGG 120
QY 431 GGTGAGCTATTCGACTACCTGTGTAAGAGGAGAGTCAAGCCCAAGAGGCGCGAAG 490
Db 121 GGTGAGCTATTCGACTACCTGTGTAAGAGGAGAGTCAAGCCCAAGAGGCGCGAAG 180
QY 491 TTCTTCCCGCAGATGTGTCTGCGCTGAGACTTCTGCAACAGCTACTTCATCTCCACAGA 550

```

```

Db 181 TTCTTCCCGCAGATGTGTCTGCGCTGAGACTTCTGCAACAGCTACTTCATGTCACAG 240
QY 551 GACCTAAAGCCCGAGAACCTGCTTTTGTGATGAGAAAACATCCGATTCAGACTTC 610
Db 241 GATCTGAAACCTGAAACCTCTGCTGTGAGCAGGAAGAACAAATCCGATTCAGACTTC 300
QY 611 GGCATGAGCTCTCTGAGAGTGGGAGCAGACCTCTCTGAGACCAAGTCTGGGCTCC 670
Db 301 GGCATGAGCTCTCTGAGAGTGGGAGCAGACCTCTCTGAGACCAAGTCTGGGCTCC 360
QY 671 TATGCTGTCCAGAGATGATTAAGGGGAGAAAATATGATGCGCGGCGAGACATGAG 730
Db 361 TATGCTGTCCAGAGATGATTAAGGGGAGAAAATATGATGCGCGGCGAGACATGAG 420
QY 731 AGCTGTGAGTCAATCTCTGCTGCTGTGAGGAGCTCTGCTGATGATGATGATGATG 790
Db 421 AGCTGTGAGTCAATCTCTGCTGCTGTGAGGAGCTCTGCTGATGATGATGATGATG 480
QY 791 CTCCGCGAGCTCTGAGAGAGTGAAGAGTGAAGGAGGAGGAGGAGGAGGAGGAG 850
Db 481 TTGCGACAGCTGTGAGAGAGTGAAGAGTGAAGGAGGAGGAGGAGGAGGAGGAG 540
QY 851 CCAGATTCGAGAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 910
Db 541 CCAGATTCGAGAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 911 CTGAGCAAAATTCAGAAACATCTTGTGATACCTAGGCGGAGGAGGAGGAGGAGGAG 970
Db 601 CTGAGCAAAATTCAGAAACATCTTGTGATACCTAGGCGGAGGAGGAGGAGGAGGAG 656
QY 971 CTGAGCAAAATTCAGAAACATCTTGTGATACCTAGGCGGAGGAGGAGGAGGAGGAG 1030
Db 657 --AGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 714
QY 1031 GACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1090
Db 715 GACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 774
QY 1091 CATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1150
Db 775 CTGAGAGAGCTGTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 834
QY 1151 GATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1210
Db 835 GATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 894
QY 1211 GACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1270
Db 895 GACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 954
QY 1271 GAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1330
Db 955 GAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1005
QY 1331 CCCAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1390
Db 1006 CCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1065
QY 1391 GCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1450
Db 1066 GCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1107
QY 1451 TCACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1510
Db 1108 ----- 1107
QY 1511 CAGAGCTGCTTCTGCGGCGCCAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1570
Db 1108 -----GTGAGCCCTTCACCCCTCAGC 1127
QY 1571 AGTGCAGCTTCACACCCCTTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1630

```


QY 255 GTCCAGAGGTCGCCATCAGATGCTGAACCGGGAGAGAGCTGCGAGTGGTGTGATGA 314
 DB 61 GTCCAGAGGTCGCCATCAGATGCTGAACCGGGAGAGAGCTGCGAGTGGTGTGATGA 120
 QY 315 AGGTGAGGCGGAGATGCGATCGTGAAGCTCAGAACACCGATGCTCCAGCTCC 374
 DB 121 AGGTGAGGCGGAGATGCGATCGTGAAGCTCAGAACACCGATGCTCCAGCTCC 180
 QY 375 AGCAGCTACGAGAACAAATATTTGTAACCTGTTGAGACACGCTCGGGGGGTG 434
 DB 181 AGCAGCTACGAGAACAAATATTTGTAACCTGTTGAGACACGCTCGGGGGGTG 240
 QY 435 AGCTATTCGACTACCTGCT 453
 DB 241 AGCTATTCGACTACCTGCT 259

RESULT 9

US-09-919-585-4
 ; Sequence 4, Application US/09919585
 ; Patent No. US20020115167A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sun, Tian-Qiang
 ; APPLICANT: Feng, Jia-Jia
 ; APPLICANT: Reinhard, Christoph
 ; APPLICANT: Fanli, Wendy J.
 ; APPLICANT: Williams, Lewis T.
 ; TITLE OF INVENTION: ISOLATION OF DROSOPHILA AND HUMAN POLYNUCLEOTIDES ENCODING PAR-1
 ; TITLE OF INVENTION: POLYPEPTIDES ENCODED BY THE POLYNUCLEOTIDES AND METHODS UTILIZI
 ; FILE REFERENCE: PP-016093.002/200130.525
 ; CURRENT APPLICATION NUMBER: US/09/919,585
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 2112
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-919-585-4

Query Match Best Local Similarity 56.4%; Score 240.6; DB 10; Length 2112;

Matches 472; Conservative 0; Mismatches 359; Indels 6; Gaps 1;

QY 177 TGGGCCCCCTATCGCTGAGAGAGAGCGCTGGGCAAGAGACAGAGCGGCTGTTAAACTCG 236
 DB 68 TTGGAACCTACCGGCTCTCTCAAGACCATTTGGCAAGGTAATTTGCCAAGTGAAGTTGG 127
 QY 237 GGGTCCACTGTCATCAGGGGTCAAGAGGTGCGCATCAAGATCGTGAACCGGAGAACTGT 296
 DB 128 CCCGACACATCTGACTGGGAGAGAGTACGTGAGATCATTTGACAAAGTCAACTGA 187
 QY 297 CGGAGTCGGGCTGATGAGAGGTGAGGAGAGATCCGATCTTAAGTCTATCAACACC 356
 DB 188 ACTCTCCACGCTCCAGAACTATTCGCGAAGTAAAGTAATTAAGTTTGAATCTATC 247
 QY 357 CACATGCTCTCAAGCTCCAGAGCTGATGAGAGAGCAAGAAATATTTGACTGCTGTTGCG 416
 DB 248 CCAACATAGTTAATTTATTTGAAGTATTTGAGACTGAGAAAAACCTCTACTTGTCTATG 307
 QY 417 AGCAGCTCGGGGGGTGAGCTATTTGACTGATGCTGTTAAAGAGAGGAGAGTGAACGCCA 476
 DB 308 AGTACGCTAGTGGCGGAGAGATTTGATTAACCTAGTGGCTCATGGCAGATGAAGAAA 367
 QY 477 AGGAGCCCGGAAGTCTTCCGCCAGATTTGTGTGGGCTGAGATTTTGGCAGACACTACT 536
 DB 368 AAGAGGCTGAGCCAAATTCGCGAGATGATGCTGTGCTGAGTACTGTCTCAGAGAAGT 427
 QY 537 CCATCGCCACAGAGACCTTAAGCCGAGAACCTGCTTTGGATGAGAAAAACAACATCC 596
 DB 428 TTATTTCCATAGAGACTTAAAGGAGAAAAACCTGCTCTTGATGCTGATATGAACATCA 487

QY 597 GCATTGACAGACTTCGCGATGCGCTCCCTGACAGGTGGGGGACAGGCTCTGAGACAGCT 656
 DB 488 AGATTGACAGACTTCGCGATGCGCTCCCTGACAGGTGGGGGACAGGCTCTGAGACAGCT 547
 QY 657 GCGGTCCTCCCTATTTGCTGTCCAGAGTGAATTAAGGGGGAATAATATGATGCGGCC 716
 DB 548 GTGGAGTCCCTCTTATGCTGCGCCCAAACTCTCCAGGGGCAAAAATAATATGATGAGCCG 607
 QY 717 GGGCAGACATGTTGAGACTGAGATCATCTCTTCCGCTGCTGCTGCTGCTGCTGCTGCT 776
 DB 608 AGTGTATGTTGAGACTGAGATCATCTCTTATGAGACTGAGACTGAGACTGAGACTGAG 667
 QY 777 TTGATGACAGCAACTCCCGCAGCTGCTGAGAGAGTGAACCGGGGCTTCCACATGC 836
 DB 668 TTGATGACAGCAACTCCCGCAGCTGCTGAGAGAGTGAACCGGGGCTTCCACATGC 727
 QY 837 CCCACTTCAATCTCCAGATTTGCCAGAGCTCTCTGAGGGGAATGATGAAGTGAACGCCG 896
 DB 728 CATTTCAATCTCCAGATTTGCCAGAGCTCTCTGAGGGGAATGATGAAGTGAACGCCG 787
 QY 897 AAAAAAGCTCAGTCTGAGAGCAATTCAGAAACATCTCTGTAACCTAG-----GCGGGA 950
 DB 788 GCAAGAGAGGACTTTAGAGCAATTCAGAAACATCTGATGATGATGATGATGATGATG 847
 QY 951 AACAGAGGACAGCCGCTGCTGAGAGCAAGCCCTGCGCCGCGGTAACCCATGCGGA 1007
 DB 848 ATGATGACACTTAAGCTTACGTGAGAGCACTCCCTGATCAAGAGACCCCGCGGGA 904

RESULT 10

US-09-919-585-7
 ; Sequence 7, Application US/09919585
 ; Patent No. US20020115167A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sun, Tian-Qiang
 ; APPLICANT: Feng, Jia-Jia
 ; APPLICANT: Reinhard, Christoph
 ; APPLICANT: Fanli, Wendy J.
 ; APPLICANT: Williams, Lewis T.
 ; TITLE OF INVENTION: ISOLATION OF DROSOPHILA AND HUMAN POLYNUCLEOTIDES ENCODING PAR-1
 ; TITLE OF INVENTION: POLYPEPTIDES ENCODED BY THE POLYNUCLEOTIDES AND METHODS UTILI
 ; FILE REFERENCE: PP-016093.002/200130.525
 ; CURRENT APPLICATION NUMBER: US/09/919,585
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 7
 ; LENGTH: 2222
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-919-585-7

Query Match Best Local Similarity 56.4%; Score 240.6; DB 10; Length 2222;

Matches 472; Conservative 0; Mismatches 359; Indels 6; Gaps 1;

QY 177 TGGGCCCCCTATCGCTGAGAGAGAGCGCTGGGCAAGAGAGAGCGGCTGTTAAACTCG 236
 DB 196 TTGGAACCTACCGGCTCTCTCAAGACCATTTGGCAAGGTAATTTGGCCAAAGTGAAGTTGG 255
 QY 237 GGGTCCACTGTCATCAGGGGTCAAGAGGTGCGCATCAAGATCGTGAACCGGAGAACTGT 296
 DB 256 CCCGACACATCTGACTGGGAGAGAGTACGTGAGATCATTTGACAAAGTCAACTGA 315
 QY 297 CGGAGTCGGTCTGATGAGAGGTGAGCGGAGATTCGCCATCTGAAAGTCAAGACACC 356
 DB 316 ACTCTCCACGCTCCAGAACTATTCGCGAAGTAAAGTAATTAAGTTTGAATCTATC 375
 QY 357 CACATGCTCTCAAGCTCCAGAGCTGATGAGAGAGCAAGAAATTTGACTGCTGTTGCG 416
 DB 376 CCAACATAGTTAATTTATTTGAAGTATTTGAGACTGAGAAAAAGCTCTACTTGTCTATG 435

QY 417 AGCAGCTCTGGGGGAGAGTATTCGACTACCTGGTAAAGAGGGAGACTGAGCCCA 476
 Db 436 AGAAGCTAGTGGGAGAGAGTATTTGATTAAGTACAGTGCATGCGAGAGTGAAGAA 495
 QY 477 AGGAGGCCCCAAAGTTCTTCGCCAGATGTGTGGCGCTGGACTTGGCCAGACTACT 536
 Db 496 AAGAGGCTCGAAGCAATTCGCCAGATAGTGTCTGCTGTGAGTACTGTCCAGAACT 555
 QY 537 CCATTCGCCAGAGACTTAAGCCCGAGAACCTGCTTTGGATGAGAAAACAACATCC 596
 Db 556 TTAATGTCATAGAGCTTAAGAGCAGAAAACCTGCTTGGATGCTGATGAAACATCA 615
 QY 597 GCATGAGACTTGGGATGGCCCTCCAGAGTGGGAGACGCTCCTGGAGACGCT 656
 Db 616 AGATTGGAGACTTGGCTTACGAATGATTCACCTTTGGAGAACGCTGGACACTTCT 675
 QY 657 GCGGATCCCCCATTTATGCGTGCAGAGTGTATTAAGGGGAAAATATGATGGCCGC 716
 Db 676 GTGGCAGTCCCCCTTATGCTGCGCCAGACTCTTCAGGGGCAAAAATATGATGACCCG 735
 QY 717 GGGCAGACATGTGAGCTGTGAGTCACTCTTCGCCCCCTGCTGGGCTGCTGCT 776
 Db 736 AGGTGATGTGTGAGGCTAGAGATTAATCTCTATACACTGTGTCAGCGGATCCGCTT 795
 QY 777 TTGATGACGACAACTCCGCCAGCTGCTGGAGAAAGTGAAGGGGCTCTTCCACATGC 836
 Db 796 TTGATGACGACAACTCCGCCAGCTGCTGGAGAAAGTGAAGGGGCTCTTCCACATGC 855
 QY 837 CCCACTTATTCCTCCAGATTTGCCAGAGCCCTCGAGGGGATGATGCAAGTGGAGCCG 896
 Db 856 CATTCATCATGCTCCAGGACTGTGAAAACCTGCTTAAGAAATTTCTTAAATCCCA 915
 QY 897 AAAAAAGCTCAGTCTGAGCAAAATCAGAAACCTCTTGTACCTAG-----GCGGA 950
 Db 916 GCAAGAGAGGACACTTATAGCAAAATCATGAAGATGATGATGATGATGATGATGATG 975
 QY 951 AACACGAGGAGAGCCGCTGCTGGAGAGCCGCTGGCCGCGGCTAGCATGCGGA 1007
 Db 976 ATGATGAACCTTAAGCTTACGCTGAGAGCCACTGACTACAGAGACCCCGCGCGA 1032

 RESULT 11
 US-09-938-842A-2657
 ; Sequence 2657, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SRIPI300-3
 ; CURRENT APPLICATION NUMBER: US/09/938, 842A
 ; PRIOR FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227, 866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264, 647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300, 111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 3379
 ; SEQ ID NO 2657
 ; TYPE: DNA
 ; LENGTH: 1539
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-2657

 Query Match 7.6%; Score 225.4; DB 9; Length 1539;
 Best Local Similarity 56.9%; Pred. No. 4,3e-43;
 Matches 434; Conservative 0; Mismatches 326; Indels 3; Gaps 1;
 QY 185 TATCGGCTGAGAAAGCGCTGGGCAAGAGACAGAGGCTGTTAACTGGGGTCCAC 244

Db 55 TACAACTTGGGAGAACTCTGTATTTGTTCTTTTGGTAGGGTGAAGATAGCTAGCAT 114
 QY 245 TGCATCAGGGTGTGAAGAGTCCGATCAAGATCGTGAACCGGGAGAACCTGCGA---G 301
 Db 115 GCATTCAGAGCAATTAAGGTGCTATCAAGATCCATCTCGCAAAATCAAGAACATG 174
 QY 302 TCGGTGCTGATGAAGTGTGAGCGGAGATGCCATCTCGAAGCTATCAACACCCACAT 361
 Db 175 GAGATGAGAGAAAGTGTGAGAGAGATCAAAATCTTGAACATATTATGATCTCTCAC 234
 QY 362 GTCTCAAGTCCACGAGCTGTACGAGAACAGAAATTTGTACCTGTCTGTGAGCAC 421
 Db 235 ATCATCGTCTATGAGGTATAGAGACTCCACAGATATTATCTGTATGAGATAT 294
 QY 422 GTCTGGGGGTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 481
 Db 295 GTGAACTCTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 354
 QY 482 GCGGAAAGTCTTCCGAGATTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 541
 Db 355 GCGAGAACTTTTTCAGCAGATTAATTAAGAGAGTGGAAATCTGCCATCGAAACATGGTG 414
 QY 542 TGCACAGAGACTTAAGCCCGAGAACCTGCTTTGGATGAGAAACATATCCGATTT 601
 Db 415 GTTCACAGAGACTCAAGCCCTGAACCTGCTTTGGACTTAATGATGATGATGATGAT 474
 QY 602 GCAGACTTGGGATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 661
 Db 475 GCTGATTTTGGCTGAGCAACATATGAGATGATGATGATGATGATGATGATGATGATG 534
 QY 662 TCCGCCATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
 Db 535 AGTCAAAATTTATGCGGCTCCAGAGTAAATTTGCGGCAAGTATATGCGGCCCTGAAGTA 594
 QY 722 GACATGTGAGCTGTGAGATCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 781
 Db 595 GATGTGAGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 654
 QY 782 GACGACAACTCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 841
 Db 655 GATGAAACATTTCCCACTTTTAAAGATTAAGGAGGAGATATACATTTACTAGC 714
 QY 842 TTGATTCCTCAGATTCAGAGCTCTGAGGGAATGATGATGATGATGATGATGATGATG 901
 Db 715 CATTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 774
 QY 902 AGCTCAGTGTGAGCAATTCAGAAACATCTTGTGATGATGATGATGATGATGATGATG 944
 Db 775 CGAGTAACTCTCTGAGATCCGCAACACCTTGTGCTCAAG 817

 RESULT 12
 US-09-815-915-3
 ; Sequence 3, Application US/09815915
 ; Patent No. US2002025931A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc.
 ; APPLICANT: Meyers, Rachel
 ; TITLE OF INVENTION: 3714, 16742, 23546, AND 13887 NOVEL
 ; FILE REFERENCE: 38155-20006.00
 ; CURRENT APPLICATION NUMBER: US/09/815,915
 ; PRIOR FILING DATE: 2001-03-23
 ; PRIOR APPLICATION NUMBER: US 60/191, 846
 ; PRIOR FILING DATE: 2000-03-24
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; TYPE: DNA
 ; LENGTH: 2352
 ; ORGANISM: Homo sapiens
 US-09-815-915-3

Tue Apr 22 09:15:28 2003

us-10-003-690-1.rmp

Page 12

Job time : 205.254 secs

```

Db 494 AGCTGTCACAGCCGCTCCATATGCAACAGGAGTGTAAACCTGAAACCTGCTGCG 553
OY 579 ATGAGAAAAAACATCCGATTCGATGCAAGACTTGGCATGGCTCCCTGAGTGGGGACA 638
Db 554 ACAGAGAAAGAACATCGGATGCGATGCGACTTTGGCATGGCTCCCTGAGTGGGGACA 613
OY 639 GCGTCTCGAGACAGTGGGGGTCCTCCATATGCTGTCCAGAGTGAATTAAGGGG 698
Db 614 GCGTGTGGAGACAGTGTGGTCCCTCCCTGAGTGTGCTGCTGCTGCTGCTGCTGCTG 673
OY 699 AAAAATATGATGGCGCGCGGACAGATGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 758
Db 674 AGAAGTGTGAGCGCGGAGAGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 733
OY 759 TCGTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 818
Db 734 TGGTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 793
OY 819 GGGGCGCTTCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 878
Db 794 GGGGCGCTTCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 853
OY 879 TGAATGAGTGTGAGCGCGGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 938
Db 854 TGAATGAGTGTGAGCGCGGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 913
OY 939 ACCTAGCGGGAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 998
Db 914 ATATAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 967
OY 999 CCATGGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1058
Db 968 AGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1027
OY 1059 CACTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1118
Db 1028 CACTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1087
OY 1119 ACCAAGAAAGATGATATATATATATATATATATATATATATATATATATATATATAT 1178
Db 1088 ACCAGGAGAGATGATATATATATATATATATATATATATATATATATATATATATAT 1147
OY 1179 AGGAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1238
Db 1148 AGGATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1207
OY 1239 CCATGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1298
Db 1208 CGATGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1267
OY 1299 TCACCGATGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1358
Db 1268 TGAC-----GGAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1318
OY 1359 AGACAGCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1418
Db 1319 AGACAGCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1378
OY 1419 TAAGCAGCCCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1478
Db 1379 TCAGCAGCCCGG----- 1391
OY 1479 CTCGAGCGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1538
Db 1392 ----- 1391
OY 1539 GTGGGGGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1598
Db 1392 -----GTGACCGCTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1431
OY 1599 CCCAAGGCTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1658

```

```

Db 1432 -----ACCCCAAGGGGACACTGTGTCACAGGCCAA 1462
OY 1659 GGGCAGTCCACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1718
Db 1463 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1516
OY 1719 GGGGAGCGCGCTGAGAGAGTGTCTCAATCCATCCAGCCAGCTTCTGGGCTCCCTC 1778
Db 1517 GAGGGGAGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1576
OY 1779 GCTTTCACCGCGGCAAGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1838
Db 1577 GCTTTCACCGCGGCAAGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1636
OY 1839 AGTCTCTCCCGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1898
Db 1637 AGTCTCTCCCGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1696
OY 1899 AGAACAATATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1958
Db 1697 AGGAGCAGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1756
OY 1959 TCCATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2018
Db 1757 TCCATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1816
OY 2019 GGGCAGAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2078
Db 1817 GGGCAGAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1876
OY 2079 TGACATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2138
Db 1877 TGACATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2118
OY 2139 GTGGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2198
Db 1919 AGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1978
OY 2199 TGATGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2258
Db 1979 TGATGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2038
OY 2259 TGACAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2277
Db 2039 TGACAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2057

RESULT 2
US-09-930-181-3
; Sequence 3, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 VI
; CURRENT APPLICATION NUMBER: US/09/930,181
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3364
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (482)..(2239)
US-09-930-181-3

Query Match 31.9%; Score 951; DB 4; Length 3364;
Best Local Similarity 69.9%; Pred. No. 7.9e-189;
Matches 1484; Conservative 0; Mismatches 440; Indels 199; Gaps 7;
OY 159 CCCAGACGCGCCCAATATGTGGGCGCTATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 218

```


DB 257 CGAGACGCGCATATGTTGGGCTTACCGGCTGGAGAACGCTGGGCGAGAG 316
OY 219 CAGGGCTGTTAACTCGGGTCCATGTCATCAGGGTGCAGAGGTCGATCAGATCG 278
DB 317 CAGGCTGGTGAAGCTGGGGTTCTACCTGCTACCTGCGCAGAGGTGGCATCAGATCG 376
OY 279 TGAACCGGGAAGCTGCGAGTGGTGGTGAAGAGTGGAGCGGAGATCGCATCG 338
DB 377 TCAACCGTGAAGCTGAGAGTGGTGGTGAAGAGTGGAGCGGAGATCGGATCG 436
OY 339 TGAAGCTCATGAGACACCATGCTCTCAAGCTCAGACGATGTCAGAGACAGAAAT 398
DB 437 TGAAGCTCATGAGACACCATGCTCTCAAGCTCAGACGATGTCAGAGACAGAAAT 496
OY 399 ATTTGAGTACCTGGTGGTGAAGAGTGGTGGTGAAGAGTGGAGCGGAGATCGGATCG 454
DB 497 ATTTGAGTACCTGGTGGTGAAGAGTGGTGGTGAAGAGTGGAGCGGAGATCGGATCG 556
OY 455 AGAAGGGGAGACTGACCGCCAGAGAGCGGAGAGTTCCTCCGAGATGTCGTCG 514
DB 557 AAGAGGGGAGGCTGAGCGCTTGAAGAGGCTGGAAGTTCCTCCGAGATGTCGTCG 616
OY 515 CTGAGCTTTCGACACGCTTCCATGCTGACAGAGACCTAAGCCGAGAACCTGCT 574
DB 617 CTGAGCTTTCGACACGCTTCCATGCTGACAGAGATCTGAAACCTGAAACCTCTG 676
OY 575 TTGAGTGAAGAAACAACATCCGATTCGACACTTCGCGATGGCTCCGTCGAGTGGG 634
DB 677 CTGAGAGAGAAACAACATCCGATTCGACACTTCGCGATGGCTCCGTCGAGTGGG 736
OY 635 GACAGCTCTGAGAGACGACTGGGGTCCCGCATATGCTGTCGAGAGTGAATTAG 694
DB 737 GACAGCTCTGAGAGACGACTGGGGTCCCGCATATGCTGTCGAGAGTGAATTAG 796
OY 695 GGGGAAATATGATGAGCGCGCGGAGACATGAGAGTGGAGTGGATCTCTGCGC 754
DB 797 GGGGAAATATGATGAGCGCGCGGAGACATGAGAGTGGAGTGGATCTCTGCGC 856
OY 755 CTGCTGTTGGAGAGACGCTGGGGTCCCGCATATGCTGTCGAGAGTGAATTAG 814
DB 857 TTGCTGTTGGAGAGACGCTGGGGTCCCGCATATGAGCAACTGTCGAGAGTGG 916
OY 815 AAGCGGGGCTCTTCCATGCTGCGGATCTCTGCGAGTGGAGAGTGGATCTCTGCGC 874
DB 917 AAGCGGGGCTCTTCCATGCTGCGGATCTCTGCGAGTGGAGAGTGGATCTCTGCGC 976
OY 875 GATGATGATGAGAGGAGCGGAGAGAGAGTGGAGAGTGGAGAGTGGATCTCTGCGC 934
DB 977 GATGATGATGAGAGGAGCGGAGAGAGAGTGGAGAGTGGAGAGTGGATCTCTGCGC 1036
OY 935 TGGTACCTGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 994
DB 1037 TGGTATATAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1090
OY 995 GTAGCCATGGGAG 1054
DB 1091 GTAGCCATGGGAG 1150
OY 1055 GATGATGATGAG 1114
DB 1151 GATGATGATGAG 1210
OY 1115 GAG 1174
DB 1211 GAG 1270
OY 1175 TGTGAG 1234
DB 1271 CAG 1330
OY 1235 TGTGAG 1294

DB 1331 TCCCGATGCTGAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1390
OY 1295 AGCATACCGAGTCCCGGGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1354
DB 1391 AGCGTAC-----GGAGGGGGGCTCCCGGGTGGTGGTGGTGGTGGTGGTGG 1441
OY 1355 GCGGAG 1414
DB 1442 GCGGAG 1501
OY 1415 CCGTAAAG 1474
DB 1502 CCGTAAAG 1514
OY 1475 GAGGTCGAG 1534
DB 1515 ----- 1514
OY 1535 AGGGTGGGGGCGCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1594
DB 1515 -----CCCGGTTGACCCCTACCCCTACAGAGAGAGAGAGAGAGAGAGAG 1558
OY 1595 GGGCCCGAGAGTCCCGGCTCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1654
DB 1559 -----ACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1585
OY 1655 CCGCGGAGAGTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1714
DB 1586 CCAAG 1639
OY 1715 GTCGGGAG 1774
DB 1640 GTCGGGAG 1699
OY 1775 CCGTAAAG 1834
DB 1700 CCGCGGAG 1759
OY 1835 CCAAGTCTCCCGGAG 1894
DB 1760 CCAAGTCTCCCGGAG 1819
OY 1895 AAG 1954
DB 1820 AAG 1879
OY 1955 ATGCTCATGCTCTTCTGTCGATCCCGAGTCCGAGTCCGAGTCCGAGTCCGAG 2014
DB 1880 ATGCTCATGCTCTTCTGTCGATCCCGAGTCCGAGTCCGAGTCCGAGTCCGAG 1939
OY 2015 TTGAGGCGGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2074
DB 1940 TTGAGGCGGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1999
OY 2075 CAGGTGATATACGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2134
DB 2000 CAGGTGATATACGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2041
OY 2135 GGTGAG 2194
DB 2042 GAG 2101
OY 2195 GAG 2254
DB 2102 AGGAG 2161
OY 2255 GCGGTCGAG 2277
DB 2162 CAGGTGATATACGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2184

RESULT 3
US-09-101-146-44


```

; LENGTH: 2652
; TYPE: DNA
; ORGANISM: Rat
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(1747)
; OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -
; OTHER INFORMATION: fragment begins at nucleotide 24 and ends with
; OTHER INFORMATION: nucleotide 1765
US-08-557-006C-39

```

```

Query Match      6.7%; Score 198.6; DB 4; Length 2652;
Best Local Similarity 54.6%; Pred. No. 2.2e-32;
Matches 418; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

```

```

OY 177 TGGGCCCCCTATCGGCTGAGAGAGCGCTGGGCAAGAGACAGAGCGGTGTTAACTCG 236
DB 45 TCGGACACTACGTGCTGGGGGACACCCCTGGGCGTGGCAGCTTGGCAAGGAAATG 104
OY 237 GGGTCCACTGCATACGCGGTGAGAGAGTCCGATCAAGATCGTGAACCGGAGAGCTGT 296
DB 105 GAAACATCAATGATGAGGAGGCAATTAAGTGAAGTAAATGATTAATGACAGAAATTC 164
OY 297 CGGAGT--CGTGTGATGATGAGGTGAGCGGAGATCGCATCTGAACTCATCGAAC 353
DB 165 GCAGTTAGATGTGTGTGAAAAATTAACGAGAAATCAAAATCTTAACTCTTTCGTC 224
OY 354 ACCACATGTCCTCAAGCTCCAGAGCTCTACGAGAACAGAAATATTGTACCTGGGTTTC 413
DB 225 ATCTCTATTAATTAACACTTACCAAGATGATGACACTCCACAGACTTTTATGATGA 284
OY 414 TGGAGCAGCTCTCGGGGGGTGAGCTATTCGACTACCTGTGTAAAGAGGAGAGACTGAGC 473
DB 285 TGGAAATATGTGTGAGGAGGAAATTTGTTGACTCATGATCTGAACAGGAGGAGTTGAAG 344
OY 474 CCAAGAGGCCCCAAATTTCTCCGCAAGATTTGTCTGCGCTGAGCTTTCGCCAGCT 533
DB 345 AGGTGGAAGCTCGCGGCTCTTCAGAGATTTGTCTGCGTGGAGTACTGTCTACAGGC 404
OY 534 ACTCATCTGCTCAGAGAGAGCTTAAGCCCGGAGACCTCTTTGGATGAGAAACAACA 593
DB 405 ACATGTTGTCCACAGGAGACTGAAGCCAGAGAACGTGTCTGTGAGCCGCCAGATGAATG 464
OY 594 TCCGATTTGACAGATTCGCGATGCGCTCCCTGCAAGTGGGAGAACCTCTCTGAGACCA 653
DB 465 CTAAAGATGCTACTGCTGACTCTCTAATATGATGTGATGATGTAATTTCTACAGACTA 524
OY 654 GCTGCGGCTCCGCCATTTATCCGCTGTCAGAGTGAATTAAGGGGAAAAATATGATGCC 713
DB 525 GCTGTGATGCGCAAAATTAATGACAGCCGAGGATCTATCAAGAAAGCTGTATGCGGGTC 584
OY 714 GCGGAGCAGATGTGAGTGTGAGTGCATCTCTTCGCGCTGCTGGGGGCTGTC 773
DB 585 CTGAGTGTGATGTGAGAGCTGTGTGTATCTCTGTATGCCCTTCTGTGACACCTTC 644
OY 774 CTTTGTATGACAGCAACCTCCGCACTGCTGAGAGAGGTGAAGCGGGGCTCTTCCACA 833
DB 645 CGTTCAGATGATGACAGCTGTGCTACGCTCTTAAAGATCCGAGGGGTGTGTCTACA 704
OY 834 TGGCCCATCTCATCTCCGCAAGATGCGCAGAGCTCTCTAGAGGAGATGATGAGAGGAGC 893
DB 705 TCCCGGAGTATCTCAACCGTTCTATTTGCCACTGTGCTGATCAACATGCTGCGAGTGAAC 764
OY 894 CCGAAAAAAGAGCTCAGTCTGAGCAAAATTCAGAAACATCTTGT 938
DB 765 CATTGAAGCAGCACTATCAAGACATACGAGACATGAATGTT 809

```

```

RESULT 7
US-08-557-006C-24
; Sequence 24, Application US/08557006C
; Patent No. 6258547
; GENERAL INFORMATION:

```

```

; APPLICANT: Berl, Rajindar K.
; APPLICANT: Carling, David
; APPLICANT: Forster, Robert A.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
; FILE REFERENCE: NCAP/PHM3588/UST
; CURRENT APPLICATION NUMBER: US/08/557,006C
; PRIOR FILING DATE: 1996-03-06
; PRIOR APPLICATION NUMBER: PCT/GB94/01093
; PRIOR FILING DATE: 1994-05-20
; PRIOR APPLICATION NUMBER: GB 9310489.1
; PRIOR FILING DATE: 1993-05-21
; PRIOR APPLICATION NUMBER: GB 9318010.7
; PRIOR FILING DATE: 1993-08-31
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 24
; LENGTH: 2761
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: cDNA of rat
; OTHER INFORMATION: Liver AMP protein kinase
US-08-557-006C-24

```

```

Query Match      6.7%; Score 198.6; DB 4; Length 2761;
Best Local Similarity 54.6%; Pred. No. 2.2e-32;
Matches 418; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

```

```

OY 177 TGGGCCCCCTATCGGCTGAGAGAGCGCTGGGCAAGAGACAGAGCGGTGTTAACTCG 236
DB 61 TCGGACACTACGTGCTGGGGGACACCCCTGGGCGTGGCAGCTTGGCAAGGAAATG 120
OY 237 GGGTCCACTGCATACGCGGTGAGAGAGTCCGATCAAGATCGTGAACCGGAGAGCTGT 296
DB 121 GAAACATCAATGATGAGGAGGCAATTAAGTGAAGTAAATCTTAATGACAGAAATTC 180
OY 297 CGGAGT--CGTGTGATGATGAGGTGAGCGGAGATCGCATCTGAACTCATCGAAC 353
DB 181 GCAGTTAGATGTGTGTGAAAAATTAACGAGAAATCAAAATCTTAACTCTTTCGTC 240
OY 354 ACCACATGTCCTCAAGCTCCAGAGCTCTACGAGTACGAGAACAGAAATATTGTACCTGGTTC 413
DB 241 ATCTCTATTAATTAACACTTACCAAGATGATGACACTCCACAGACTTTTATGATGA 300
OY 414 TGGAGCAGCTCTCGGGGGGTGAGCTATTCGACTACCTGTGTAAAGAGGAGAGCTGAGC 473
DB 301 TGGAAATATGTGTGAGGAGGAAATTTGTTGACTCATGATCTGAACAGGAGGAGTTGAAG 360
OY 474 CCAAGAGGCCCCAAATTTCTCCGCAAGATTTGTCTGCGCTGAGCTTTCGCCAGCT 533
DB 361 AGGTGGAAGCTCGCGGCTCTTCAGAGATTTGTCTGCGGCTGAGTACTGTACAGGC 420
OY 534 ACTCATCTGCTCAGAGAGAGCTTAAGCCCGGAGAACCTGCTTTGGATGAGAAACAACA 593
DB 421 ACATGTTGTCCACAGGAGCTGAGCAGAGAACGATGTGTCTGAGCCGCCAGATGAATG 480
OY 594 TCCGATTTGACAGATTCGCGATGCGCTCCCTGCAAGTGGGAGAGAGCTCTCTGAGACCA 653
DB 481 CTAAAGATGCTACTGCTGACTCTCTAATATGATGATGATGATGATGATTTCTAGAACTA 540
OY 654 GCTGCGGCTCCGCCATTTATGCTGTCCAGAGTGAATTAAGGGGAAAAATATATGATGCC 713
DB 541 GCTGTGATGCGCAAAATTAATGACAGCCGAGAGGATCTCATCTCAGAGAGCTGTATCGGGTC 600
OY 714 GCGGAGCAGATGTGAGAGTGTGAGATCATCTCTTCGCGCTGCTGTGGGGGCTGTC 773
DB 601 CTGAGTGTGATATCTGAGAGTGTGTGTATCTGTATGCTGTGCTGTGTCGACCTCC 660
OY 774 CTTTGTATGACAGCAACCTCCGCAAGTCTGAGAGGAGTGAAGCGGGGCTCTTCCACA 833
DB 661 CATTGACATGAGCAGCTGCTACGCTCTTAAAGAAATCCGAGGGGTGTGTTCTTACA 720
OY 834 TGGCCCATCTCATCTCTCAGATTTGCCAGAGCTCTCTGAGGGGAATGATCGAAGTGAAGC 893

```

Db 721 TCCCGAGTATCTCAACCGCTTCTATTGCGCACTGCTGATGACATGCTGGCAGTGAGCC 780
 Oy 894 CCGAAAAAGGCTCAGTCTGTGAGCAAAATTCGAAACATCTTGGT 938
 Db 781 CCTTGAAGGAGCACTATCAAGACATACAGAGCATGAATGT 825

RESULT 8

US-08-557-006C-37
 ; Sequence 37, Application US/08557006C
 ; Patent No. 6258547
 ; GENERAL INFORMATION:
 ; APPLICANT: Berl, Rajindar K.
 ; APPLICANT: Carling, David
 ; APPLICANT: Forster, Robert A.
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
 ; FILE REFERENCE: NGAP/PM37588/JST
 ; CURRENT APPLICATION NUMBER: US/08/557,006C
 ; CURRENT FILING DATE: 1996-03-06
 ; PRIOR APPLICATION NUMBER: PCT/GB94/01093
 ; PRIOR FILING DATE: 1994-05-20
 ; PRIOR APPLICATION NUMBER: GB 9310489.1
 ; PRIOR FILING DATE: 1993-05-21
 ; PRIOR APPLICATION NUMBER: GB 9318010.7
 ; PRIOR FILING DATE: 1993-08-31
 ; NUMBER OF SEQ ID NOS: 44
 ; SOFTWARE: Patentln Ver. 2.1
 ; SEQ ID NO 37
 ; LENGTH: 1736
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Rat liver AMP
 ; OTHER INFORMATION: protein kinase
 US-08-557-006C-37

Query Match 6.1%; Score 182.6; DB 4; Length 1736;
 Best Local Similarity 53.3%; Pred. No. 4e-29;
 Matches 408; Conservative 0; Mismatches 354; Indels 3; Gaps 1;

Oy 177 TGGGCCCTTCGCGTGGAGAGAGCGTGGCAAGACAGACAGGGGCTTAACTCG 236
 Db 38 TCGGACACTAGTGTGCTGGGACACGCTGGGCTGGCAGCTTGGCAAGAGATTG 97
 Oy 237 GGGTCCACTGCATCAGCGGTGAGAGGTCCGATCAAGATGCTGAACCGGAGAGCTGT 296
 Db 98 GAAACATCAATTAAGCGCCATTAAGTGGCAGTTAAATCTTAATAGACAGAAATTG 157
 Oy 297 CGGAGT---CGGTCTGATGAAGGTGAGGCGGAGATCGCATCTGAGCTCATCGAAC 353
 Db 158 GCAGTTAGATGTGTGGAAAAATAAAGCAGAAATCAAAATCTAAACCTTTGCTC 217
 Oy 354 ACCACATGCTCAAGCTCCAGAGCTCTACGAGAACAAAGAAATTTGTACCGGTTG 413
 Db 218 ATCTCTATTAATTAACACTATACAGGTGATCAGCATCCAAAGATTTTATGGTAA 277
 Oy 414 TGGAGCAGCTCTCGGGGGTGAAGTATTCGACTACCTGGTAAAGAGGAGAGCTGAGCG 473
 Db 278 TGAATATGTGTGAGAGTGAATTTATTTGACTACATCTGAAGATGAGAGGGTGAAG 337
 Oy 474 CCAAGAGGCGCCGAAATCTTCCGCAAGATGTGTCTGGCGTGAATCTTGCACAGCT 533
 Db 338 AGATGGAAGCAGGCGCTCTTTCAGCAGATTCGTCTGTGATTAATCTCATAGGC 397
 Oy 534 ACTCCATCTCCACAGAGACTTAAGCCGGAAGCTGCTTTGGATGAGAAAACAACA 593
 Db 398 ATATGTTGTTCACTGAGACCTGAAACAGAGAAATGCTCTGTGGATCAGACATGAATG 457
 Oy 594 TCCGATTTGAGAGCTTGGAGATGGCGTCCCTGACAGTGGGGGAGACAGCTCCCTGAGACA 653
 Db 458 CCAAGATAGCGGATTTGGGATTAATCTAATAGATGTGAGATGGAATTTTTCGAAACGA 517

Oy 654 GCTCGGGTCCCCCATTAATGCGTGTCCAGAGTGAATTAAGGGGAAAAATATGAGCC 713
 Db 518 GTTCCGATCTCCAAATTAATGACAGACCTGAAATCATCTCAGAGATTTATGAGCGTC 577
 Oy 714 GCCGGCAGCATGTGGAGGTGGAGTCAATCCCTTCGCCCTGCTGGGGGCTGCG 773
 Db 578 CTGAAGTTGATCTGGAGCTGAGTGTGTCTGTATGCTCTCTTTGTTGGACCTCC 637
 Oy 774 CTTTATGACAGCAACCTCCGCGAGCTGTGAGAGGTGAAGCGGGCGCTTCCACA 833
 Db 638 CATTGTATGATGAGCATGTAACGTACGTATTTAAGAGATCCGAGGGGCTGCTTTATA 697
 Oy 834 TCGCCCACTTCAATCTCCGAGATTGCGAGACCTCTCGAGGGGATATGATCAAGTGAAC 893
 Db 698 TCCAGATATCTCAATCTGCTGTCCGCACTCTCTGATGATGATGCTGCGAGTTGACC 757
 Oy 894 CCGAAAAAGGCTCAGTCTGAGCAAAATTCGAAACATCTTGGT 938
 Db 758 CACTGAAGGAGCACTATCAAGACATTAAGAGACATGAATGT 802

RESULT 9

US-08-557-006C-36
 ; Sequence 36, Application US/08557006C
 ; Patent No. 6258547
 ; GENERAL INFORMATION:
 ; APPLICANT: Berl, Rajindar K.
 ; APPLICANT: Carling, David
 ; APPLICANT: Forster, Robert A.
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
 ; FILE REFERENCE: NGAP/PM37588/JST
 ; CURRENT APPLICATION NUMBER: US/08/557,006C
 ; CURRENT FILING DATE: 1996-03-06
 ; PRIOR APPLICATION NUMBER: PCT/GB94/01093
 ; PRIOR FILING DATE: 1994-05-20
 ; PRIOR APPLICATION NUMBER: GB 9310489.1
 ; PRIOR FILING DATE: 1993-05-21
 ; PRIOR APPLICATION NUMBER: GB 9318010.7
 ; PRIOR FILING DATE: 1993-08-31
 ; NUMBER OF SEQ ID NOS: 44
 ; SOFTWARE: Patentln Ver. 2.1
 ; SEQ ID NO 36
 ; LENGTH: 1783
 ; TYPE: DNA
 ; ORGANISM: Human AMP protein kinase
 US-08-557-006C-36

Query Match 6.1%; Score 182.6; DB 4; Length 1783;
 Best Local Similarity 53.3%; Pred. No. 4.1e-29;
 Matches 408; Conservative 0; Mismatches 354; Indels 3; Gaps 1;

Oy 177 TGGGCCCTTCGCGTGGAGAGAGCGTGGCAAGACAGACAGGGGCTTAACTCG 236
 Db 39 TCGGACACTAGTGTGCTGGGACACGCTGGGCTGGCAGCTTGGCAAGATTG 98
 Oy 237 GGGTCCACTGCATCAGCGGTGAGAGGTCCGATCAAGATGCTGAACCGGAGAGCTGT 296
 Db 99 GAAACATCAATTAAGCGCCATTAAGTGGCAGTTAAATCTTAATAGACAGAAATTG 158
 Oy 297 CGGAGT---CGGTCTGATGAAGGTGAGGCGGAGATCGCATCTGAGCTCATCGAAC 353
 Db 159 GCAGTTAGATGTGTGGAAAAATAAAGCAGAAATCAAAATCTAAACCTTTGCTC 218
 Oy 354 ACCACATGCTCAAGCTCCAGAGCTCTACGAGAACAAAGAAATTTGTACCTGTTG 413
 Db 219 ATCTCTATTAATTAACACTATACAGGTGATCAGCATCCAAAGATTTTATGGTAA 278
 Oy 414 TGGAGCAGCTCTCGGGGGTGAAGTATTCGACTACCTGGTAAAGAGGAGAGCTGAGCG 473
 Db 279 TGAATATGTGTGAGAGTGAATTTATTTGACTACATCTGAAGATGAGAGGGTTGAAG 338
 Oy 474 CCAAGAGGCGCCGAAATTTTCCGCAAGATGTGTCTGGCGTGAATCTTGCACAGCT 533

TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
FILE REFERENCE: NSAP/PM75588/US
CURRENT APPLICATION NUMBER: US/08/557,006C
CURRENT FILING DATE: 1996-03-06
PRIOR APPLICATION NUMBER: PCT/GB94/01093
PRIOR FILING DATE: 1994-05-20
PRIOR APPLICATION NUMBER: GB 9310489.1
PRIOR FILING DATE: 1993-05-21
PRIOR APPLICATION NUMBER: GB 9318010.7
PRIOR FILING DATE: 1993-08-31
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 44
LENGTH: 1747
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: gene
LOCATION: (1)..(1747)
OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -
OTHER INFORMATION: fragment begins at nucleotide 24 and ends with
US-08-557-006C-44

Query Match 5.9%; Score 175.2; DB 4; Length 1747;
Best Local Similarity 54.4%; Pred. No. 1.4e-27;

Matches 419; Conservative 0; Mismatches 343; Indels 8; Gaps 3;

QY 177 TGGGCCCTATCGCTGAGAGAGAGCTGGGCAAGAGACAGAGCGCTGGTAACTCG 236
DB 38 TCGGACACTGCTGCTGGGGAGACACCTGGGCGCTGGGCAAGAGATG 97
QY 237 GGGTCACTGACATCAGGGGTGAGAGTCCGCAATCAAGAGTGTGAACCGGAGAGCTGT 296
DB 98 GAAACATCAATGAGAGAGAGAGAGAGTGAAGATCTTAAATGAGAGAGATTC 157
QY 297 CGGAGT---CGGTGATGATGAGGTGAGAGAGAGATCGCATCTGTAAGTCAATCGAAC 353
DB 158 GCAGTTAGATGTTGTTGGAAAAATTAACGAGAAATTCAAATCTTAACCTTTCGTC 217
QY 354 ACCCATGTCCTCAAGCTCCAGAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 413
DB 218 ATCCATATATATCAAACTTACCAAGATGATGATGATGATGATGATGATGATGATGATG 277
QY 414 TGGAGACGCTCTGGGGGTGAGAGTATTCGACTACCTGCTTAAAGAGAGAGAGAGAGAG 473
DB 278 TGGAAATGCTGCTGAGAGTGAATGTTGCTGATGATGATGATGATGATGATGATGATG 337
QY 474 CCAAG 531
DB 338 AGGTGAG 397
QY 532 CTAATCATCTGACAG 591
DB 398 GCACATGCTGCTGAG 457
QY 592 CAGCCGATGAG 651
DB 458 TGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 517
QY 652 CAGCTCGGGTCCCGCATATATGCTGATGATGATGATGATGATGATGATGATGATGATG 711
DB 518 TACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 577
QY 712 CCGCCGAG 771
DB 578 TCTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 637
QY 772 GCGCTTGAAGAG---GACAACTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 828
DB 638 CCGCTGAG 697
QY 829 CCAGATGCCCACTTATCTCTCCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 888

DB 698 CTACATCCCGAGATATCTCAACCTTCTATGCTGCTGATGACATGCTGACAGT 757
QY 889 GGAGCCCGAAG 938
DB 758 GGAGCCCGTGAAG 807

RESULT 12

US-09-930-181-5
Sequence 5, Application US/09930181
Patent No. 6455292

GENERAL INFORMATION:

APPLICANT: Origene Technologies

TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas

FILE REFERENCE: 16U 101 VI

CURRENT APPLICATION NUMBER: US/09/930,181

NUMBER OF SEQ ID NOS: 18

SOFTWARE: Patent In version 3.0

SEQ ID NO 5

LENGTH: 213

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(213)

US-09-930-181-5

Query Match 4.9%; Score 145; DB 4; Length 213;
Best Local Similarity 86.5%; Pred. No. 1.4e-21;

Matches 160; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 159 CCCAGCAGCCCAATATGTTGGGCCCTATCGGCTGAGAGAGAGAGAGAGAGAGAGAG 218
DB 29 CCGAGCAGCGGAGAGATGTTGGGCCCTACCGGCTGAGAGAGAGAGAGAGAGAGAGAG 88
QY 219 CAGGCTGCTTAATCTGGGGTCCATGCTGATGATGATGATGATGATGATGATGATG 278
DB 89 CAGGCTGCTGAAGCTGGGGGTGATGATGATGATGATGATGATGATGATGATGATG 148
QY 279 TGAACCGGAGAGAGCTGCTGAGTGGGTGCTGATGATGATGATGATGATGATGATG 338
DB 149 TCAACCGTGAAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 208
QY 339 TGAAG 343
DB 209 TGAAG 213

RESULT 13

US-09-159-385-3
Sequence 3, Application US/09159385
Patent No. 5958748

GENERAL INFORMATION:

APPLICANT: AKIRA, SHIZUO

TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE

FILE REFERENCE: PR-569

CURRENT APPLICATION NUMBER: US/09/159,385

EARLIER FILING DATE: 1997-09-23

NUMBER OF SEQ ID NOS: 8

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 3

LENGTH: 2132

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (94)..(1455)

US-09-159-385-3

Query Match	4.6%;	Score 135.6;	DB 2;	Length 2132;
Best Local Similarity	54.1%;	Pred. No. 2.6e-19;		
Matches 360; Conservative	0;	Mismatches 279;	Indels 27;	Gaps 3,

QY	163	GCAGCCCAAAATATGTGGGCCCCCTTATCGGTGGAGACAACAGCTGGGGCAAGACAGACAGG	222
Db	108	GCAGAGAGAGACTGGAGGACCACTTTATGAGANTGGGGAGAGAGACTGGGGCAGCGCCAGTTTGC	167
QY	223	GCTGTTAAACTCGGGTCCACTGCATACAGGGTGCAGAGAGGTGGCCATCAAGATGTA	282
Db	168	GATGTCGCGAATGACCGGGCAGAGAGGGCAGGGCCAGAGAGATACCAACCCAAAGTTCATCAA	227
QY	283	CCGGAGAGAGCTGTC-----GAGTCGGTGTCTATGAAAGGTGGAGCGGGAGAT	330
Db	228	GAAACGCCGCTGTCAATCAGCCGGCCTGGGGTATGCCGGAGAGATCGAGCGGGAGGT	287
QY	331	CGCCATCCTGAAAGCTCATTCGAACACCCACATATGTCCTCAAGCTCCACACAGCTTACAGGAA	390
Db	288	GAAATCCTCGGGAGATACCGGACACCCCAACATCATCCCTGCAGCAATCTTCGAGAA	347
QY	391	CAAAATAATTTGTACTGTGTTCTGGAGACAGCTCGGGGGGTAGCTATTTCACTACT	456
Db	348	CAAAACGAGACTGGTCTCATCTCATCTGGAGCTGTCTCGGGGGAGAGCTCTTGTCTCT	407
QY	451	GGTAAAGAGGGAGACTGACGCCCAAGAGAGGCCCAAGATTTCTCCGACAGATTGTGTC	510
Db	408	GGCGAGAAAGATGTCGTGAGAGAGACAGAGCCACCAAGTCTCTCAAGAGATCTGGA	467
QY	511	TGCGCTGGACTTCTGGCCACAGCTACTCCATCTGCCACAGAGACTTAAAGCCCGAAGACT	570
Db	468	CGGGTTTACACTCTGCACCTTAAGGCCATCCGACACTTTGACTGGAAGCCGGAAAAAT	527
QY	571	GCTTTTGGATAGAAAAAC-----AAATCCGCAATTGCAACATTTCCGATAGGC	618
Db	528	CATCTCTGTGACAAAGACGTGGCCCAACCCACGAATATAAGCTACTCACTTCGGCATCGC	587
QY	619	GTCCTTCAGAGTGGGGACAGCCCTCTGTGAGACCAAGCTGGGGGTCCCCCATTAATGCGTG	678
Db	588	GCAACAAATCAGAGCGGGGGAACGATTCAAGAACATCTTGGGCAACCCGGAGTTTGTGGC	647
QY	679	TCCAGAGGTGATTAAGGGGGGAAAAAATATGATAGCGCGCGGGGCAACATATGGAGCTGTGG	738
Db	648	CCCAGAGATTGTGAACATAATGAGCCGCTG---GGCTTGAAGCGGACACATGTGGACATCGG	704
QY	739	AGTCATCTCTTCGCCCCCTGCTGATGGGGGCTTGGCCCTTTGATGACGACAACATCTCGCCA	798
Db	705	TGTATCATCACTATATCTCTCTGAGCGGTGATCCCGTCTCTCTGGCGAGAACAAAGCAGGA	766
QY	799	GCTGCT 804	
Db	765	GAGCGT 770	

RESULT 14
US-09-186-277-3
; Sequence 3, Application US/09186277

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (94) .. (1455)
US-09-186-277-3

```

	Query Match	Similarity	4.66;	Score 135.6;	ID 4;	Length 2132;
	Best Local Similarity	54.13;	Pred.	No.2.6e-19;		
	Matches 360;	Conservative	0;	Mismatches 279;	Indels	27; Gaps
OY	163	GCACGCCCCAATATATGTGGGCCCTTATCGGCTTGAGNAGNCGCTGGGCAAGAAGCAGACAGG	222			
Dd	108	GCAGGAGGACGCTGGAGGACATTTATGAGATGGGGGAGGAGCGTGGGCGAGCGCCATTTCG	167			
OY	223	GCTGTTAAACTCGGGGTCCACTGCATACGGGTGACAGSTGCGCCATCAAGATCTGAA	287			
Dd	168	GATCCTGGGGAAGTCCCGGCAAGAGGCGACGGGCAAGGAGTACGACCAGCATTCATCAA	227			
OY	283	CCGGGAGAAGCTGTC-----GGAGTCGCTGCTGATGAAAGGTGGAGCGGGAGAT	330			
Dd	228	GAGGGCGCGCTGTGATTCACAGCCGCGCTGGGGGTAGCGCGGAGGAGATCGAGCGGGAGGT	287			
OY	331	CGCCATCTCTGAAGCTCATCGAACAACCCATGATCTCCTCAAGCTCCAGCGATCTACGAAA	397			
Dd	288	GAACTCTCTGGGGGAGATCCGGCACGCCAACATATCATCCTCGACGACATCTTCGAAA	347			
OY	391	CAGAAATATTTGTACTGGTTCTTGAGACGACGCTCTGGGGGGTAGCTATTCGACTACT	455			
Dd	348	CAAGCGGACGCTGGCTCATCTCTGAGACTGGTGTCTGCGCGGAGAGCTTTGACTTCT	407			
OY	451	GGTAAAGAGGGGAGACGTGAGCCCAAGSAGGCCGAAAGTCTTCCGCCGATGTGTGTC	510			
Dd	408	GCGGGAAGAGATCGCTGACGAGGAGAGAGAGGACCAACCCAGTTCTCTCAAGCGATCTTGGA	467			
OY	511	TGCGCTGACTTCTGCCACGACTACTCCATCTGGCACAGACCTTAAGCCCGAGACT	570			
Dd	468	CGGGTCTACATACTGCACTCTAAGCGATGGCACACTTTGACTGAAGCCGGAACAT	527			
OY	571	GCTTTGATGAAAAAC-----AACATCCGATTCGACACTTCGGCATGGC	618			
Dd	528	CATGCTGCTGGACAAGAACGTGCCAACCCAGATCAAGCTCATCGACTTCGGCATTCGC	587			
OY	619	GTCCCTGCAAGGGGGGACACGCTCTGAGAACCAAGCTGGGGGTCCCCTTATNGGTG	678			
Dd	588	GCACAGATTCGAGGGGGGGAAGAGTTAABAACATCTTGGCACCCCGGAGTTTGTGGC	647			
OY	679	TCCAGAGCTGATTAAGGGGGAAAAATATGATGAGCGCGCGGACACATATGTGAGCTGTGG	738			
Dd	648	CCCAAGATTTGAACTATGAGCGGCTG---CGCCTGAGGGGAGCATGTGAGACATCGG	704			
OY	739	AATCATCTCTTTCGCCCTGCTCGTGGGGGCTCTGCGCTTGATGACGACAACTTCGGCCA	798			
Dd	705	TGTCATCACTAATATCTCTCTGAGAGGATGATCCCGTCTCTGGCGGAGACCAAGCAGGA	764			
OY	799	GCTGCT 804				
Dd	765	GACGCT 770				

```

? GENERAL INFORMATION:
? APPLICANT: AKIRA, SHIUCHI
? APPLICANT: KAWAI, TARO
? TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
? FILE REFERENCE: 081356/0128
? CURRENT APPLICATION NUMBER: US/09/186,277
? CURRENT FILING DATE: 1998-11-05
? EARLIER APPLICATION NUMBER: JP97/261589
? EARLIER FILING DATE: 1997-09-26
? NUMBER OF SEQ ID NOS: 8
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 3
? LENGTH: 2132
? TYPE: DNA
? ORGANISM: Homo sapiens

```

SEQ ID NO 4
LENGTH: 1429
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (10)..(1353)
US-09-159-385-4

Query Match 4.4%; Score 130.6; DB 2; Length 1429;
Best Local Similarity 56.5%; Pred. No. 2.5e-18;
Matches 291; Conservative 0; Mismatches 209; Indels 15; Gaps 2;

QY 308 CAGATGAAGTGGAGGAGATGCGCATCTGAAAGCTCAGAACCCACATGTCCTC 367
DB 181 CGGAGAGATCGAACCGAGGTAGCATCTCGCGAGATCCGCCACCCCAACATCATA 240
QY 368 AAGCTCCAGCAGCTCTCTACAGAACAAATATTTGACTGTTCTGGAGCAGCTCG 427
DB 241 ACACCTGATGACGTTCTCGAAGAACAGACATGTTGTTCTGATCTGAGCTGCTCC 300
QY 428 GGGGGTGAGCTATTCGACTACCTGGTAAAGAGGGAGACTGACGCCCAAGAGGCCGA 487
DB 301 GGTGGCGAGCTTTTCGACTTCCTGGCCGAGAGAGTCACTGACGAGATGAGGCCAAG 360
QY 488 AAGTCTCCGCGAGATGTGTCTGCGCTGAGACTTCTGCCACAGCTACTCCATCTGCCAC 547
DB 361 CAGTTCCTCAACAAATCCTAGACGAGTGTCTCCTACCTGACCTCCAGCGCATCGCACAC 420
QY 548 AGAGACCTAAAGCCCGAAGACCTGCTTTGGATGAGAAAACAACATCCGC----- 598
DB 421 TTGTACCTGAAGCCCGAAGACATCATGTGCTGACAAAGCACGCGACGCCGCCGCAATT 480
QY 599 ---ATTGCAGACTTCGGCATGGCGTCCCTGAGGTGGGGGACAGCCTCTGGAGACACAGC 655
DB 481 AAGCTCATGCACTTTGGCATGCGCACAGATCGAGGCTGGCAGGAGTTCAAGAACATC 540
QY 656 TGGGGTCCCGCCATTTGGCGTTCAGAGGTATTAAGGGGGAATAATGATGAGCCGC 715
DB 541 TTTGGCACACCCGAGTTTGGCGTCCCGAGATCGTGAACATGAGCCA---CTTGGCTTG 597
QY 716 CGGCGAGACATGTGAGCTGTGAGTATCCTCTTCGCCCTGCTGCTGGGGCTCTGCC 775
DB 598 GAGGCTGACATGTGAGCATTTGGGCTATCAGCTACATCTCTGAGCGAGCGTCCCA 657
QY 776 TTGTATGACGACAACCTCGGCCAGCTCTGGAGAA 810
DB 658 TTCTGGGGGAGACCAAGACGAGAGCGCTGACGAA 692

Search completed: April 22, 2003, 00:25:56
Job time: 83.9763 secs

BASE COUNT	ORIGIN
207 a	246 c 274 g 160 t 2 others

Query Match	25.5%;	Score 759.8;	DB 14;	Length 889;
Best Local Similarity	98.5%;	Pred. No. 7e-95;		
Matches 788;	Conservative	0;	Mismatches 8;	Indels 4;
				Gaps 2

QY	189	GGCTGGAAAGACCCCTGGGGCAAGAGACAGACAGGGGCTGGTTAAACTCGGGGCTCCACTGCA	248
Db	9	GGGTGGAAAGACCCCTGGGGCAAGAGACAGACAGGGGCTGGTTAAACTCGGGGCTCCACTGCA	68
QY	249	TCACGGGCTCAGAAAGGTGGCCATCAGATCTGAAACGGGGAGAAAGCTGTGGAGTGGGTGC	308
Db	69	TCACGGGCTCAGAAAGGTGGCCATCAGATCTGAAACGGGGAGAAAGCTGTGGAGTGGGTGC	128
QY	309	TGATGAAGGTGAGCCGGGAGATGTGCCATCTCGAAGCTCATCGAACCACCACATGTCTCTCA	368
Db	129	TGATGAAGGTGAGCCGGGAGATGTGCCATCTCGAAGCTCATCGAACCACCACATGTCTCTCA	188
QY	369	AGCTCCACGAGCTGTACGAGAAACAAGAAATATTTGTACCTGGTCTTGAGACAGCTTCGG	428
Db	189	AGCTCCACGAGCTGTACGAGAAACAAGAAATATTTGTACCTGGTCTTGAGACAGCTTCGG	248
QY	429	GGGGTAGCTTATTGACTACTCGTAAAGAGGGGAGACTGACGCCACAGAGAGGCCGAA	488
Db	249	GGGGTAGCTTATTGACTACTCGTAAAGAGGGGAGACTGACGCCACAGAGAGGCCGAA	308
QY	489	AGTTCTTCCGGCAGATTGTGCTCGCTGGACCTCTCCACAGCTACTCCATCTGCCACA	548
Db	309	AGTTCTTCCGGCAGATTGTGCTCGCTGGACCTCTCCACAGCTACTCCATCTGCCACA	368
QY	549	GAGACCTTAAAGCCCGAGAACCTGCTTTTGGATGAGAAAAACAACATCCGATTTGCAGACT	608
Db	369	GAGACCTTAAAGCCCGAGAACCTGCTTTTGGATGAGAAAAACAACATCCGATTTGCAGACT	428
QY	609	TCGGCATGGCGCTCCCTCGTAGSTGGGGGACACGCTCTGAGAGACAGTGGGGGTCCCCC	668
Db	429	TCGGCATGGCGCTCCCTCGTAGSTGGGGGACACGCTCTGAGAGACAGTGGGGGTCCCCC	488
QY	669	ATTATGCTGTCCAGAGTGTATTAAGGGGGAAAAATATGATGGCCGCCGGGACAGACTGT	728
Db	489	ATTATGCTGTCCAGAGTGTATTAAGGGGGAAAAATATGATGGCCGCCGGGACAGACTGT	548
QY	729	GGAGCTGTGAGTATCTCTCTGCGCCCTGCTGTTGGGGGCTCTGCGCTTTGATGAGACA	788
Db	549	GGAGCTGTGAGTATCTCTCTGCGCCCTGCTGTTGGGGGCTCTGCGCTTTGATGAGACA	608
QY	789	ACCTCCGGCACCTCTGAGAGAAAGGTGAAACGGGGCGCTTCCACATGCCCATCTCAATTC	848
Db	609	ACCTCCGGCACCTCTGAGAGAAAGGTGAAACGGGGCGCTTCCACATGCCCATCTCAATTC	668
QY	849	CTCCAGATTGCGACAGGCTCTCTGAGGGGAATGATCGAATGAGAACCCGAAAAAGGCTCA	908
Db	669	CTCCAGATTGCGACAGGCTCTCTGAGGGGAATGATCGAATGAGAACCCGAAAAAGGCTCA	728
QY	909	GCTCTGGAGCAATTCAGAAACATCTCTTGG---TACCTAAGCGGGGAAACACAGAGCCAGACC	968
Db	729	GCTCTGGAGCAATTCAGAAACATCTCTTGTATCNCATAGGCGGGGAAACACAGAGCCAGACC	788
QY	966	CGTG-CCGTGGAGCAGCCCC 984	
Db	789	CGTGCCCTGGAGCAGCCCC 808	

RESULT 2	BQ717903	881 bp	mRNA	linear	EST 16-JUL-2002
LOCUS	BQ717903				
DEFINITION	BQ717903				
ACCESSION	BQ717903				
VERSION	BQ717903.1	GI:21856800			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 881)				
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				

Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LLM15580 row: c column: 11
High quality sequence stop: 620.

FEATURES

SOURCE

```

/organism="Homo sapiens"
/db_xref="taxon:3606"
/clone IMAGE:6166946"
/clone_lib="lupski_sym pathetic_trunk"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site:1:
NotI; Site:2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-GACGACCCACCGGCTCCG-3' and
5'-GACGAGTCTGATGATGACGAGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

```

Query Match	24.3%	Score 723.6	DB 14	Length 881
Best Local Similarity	95.3%	Pred. No. 6	2e-90	
Matches 755	Conservative 0	Mismatches 36	Indels 1	Gaps 1

QY	570	TGCTTTTGGATGAGAAAAACAATCCGATTCGGAGATTCGCGATCGCCTCCCTGGAG	629
QY <td>1</td> <td>TGCTTTTGGATGAGAAAAACAATCCGATTCGGAGATTCGCGATCGCCTCCCTGGAG</td> <td>60</td>	1	TGCTTTTGGATGAGAAAAACAATCCGATTCGGAGATTCGCGATCGCCTCCCTGGAG	60
QY <td>630</td> <td>TGGGGGACAGCCTCCTGGAGACACAGCTGCGGGTCCCCCATTTATGCTGTCCAGAGGTGA</td> <td>688</td>	630	TGGGGGACAGCCTCCTGGAGACACAGCTGCGGGTCCCCCATTTATGCTGTCCAGAGGTGA	688
Db <td>61</td> <td>TGGGGGACAGCCTCCTGGAGACACAGCTGCGGGTCCCCCATTTATGCTGTCCAGAGGTGA</td> <td>120</td>	61	TGGGGGACAGCCTCCTGGAGACACAGCTGCGGGTCCCCCATTTATGCTGTCCAGAGGTGA	120
QY <td>690</td> <td>TTAAGGGGAAAAATATGATGAGCGCGCGGAGACATGTGGAGCTGTGAGATCATCTCT</td> <td>749</td>	690	TTAAGGGGAAAAATATGATGAGCGCGCGGAGACATGTGGAGCTGTGAGATCATCTCT	749
Db <td>121</td> <td>TTAAGGGGAAAAATATGATGAGCGCGCGGAGACATGTGGAGCTGTGAGATCATCTCT</td> <td>180</td>	121	TTAAGGGGAAAAATATGATGAGCGCGCGGAGACATGTGGAGCTGTGAGATCATCTCT	180
QY <td>750</td> <td>TGCGCCTCTGCTGTGGGGCTGTGCCCTTTGATGTGACGACAAACCTCCGACAGTGGTGAGA</td> <td>809</td>	750	TGCGCCTCTGCTGTGGGGCTGTGCCCTTTGATGTGACGACAAACCTCCGACAGTGGTGAGA	809
Db <td>181</td> <td>TGCGCCTCTGCTGTGGGGCTGTGCCCTTTGATGTGACGACAAACCTCCGACAGTGGTGAGA</td> <td>240</td>	181	TGCGCCTCTGCTGTGGGGCTGTGCCCTTTGATGTGACGACAAACCTCCGACAGTGGTGAGA	240
QY <td>810</td> <td>AGGTGAAAAGGGGCGTCTTCACATGCCCCCATTCATTTCTCCAGATGGCCAGAGCCCTCC</td> <td>865</td>	810	AGGTGAAAAGGGGCGTCTTCACATGCCCCCATTCATTTCTCCAGATGGCCAGAGCCCTCC	865

Db 241 AGTGAACGGGGGCTTCCATGCCCCACCTTCATCTCCAGATTCCAGAGCCCTCC 300
QY 870 TGAGGGGAATGATCGAATGAGGCCGAAAAAGGCTCAGTCTGAGCAATTCAGAAAC 929
Db 301 TGAGGGGAATGATCGAATGAGGCCGAAAAAGGCTCAGTCTGAGCAATTCAGAAAC 360
QY 930 ATCTTGATCTAGAGGGGAAAAACAGAGCCAGACCCGCTGCTGAGAGCCGCTGGCC 989
Db 361 ATCTTGATCTAGAGGGGAAAAACAGAGCCAGACCCGCTGCTGAGAGCCGCTGGCC 420
QY 990 GCGGGGATGAGCAGGAGCCGAGCTGCTCAGTCCAGAGAGAGTGGAGCCGAGCTCCAGAGA 1049
Db 421 GCGGGGATGAGCAGGAGCCGAGCTGCTCAGTCCAGAGAGAGTGGAGCCGAGCTCCAGAGA 480
QY 1050 GCATGATCAGTCTGAGTCTGCTCAGAGAGCCGAGAGAGTGGAGCCGAGCTCCAGAGA 1109
Db 481 GCATGATCAGTCTGAGTCTGCTCAGAGAGCCGAGAGAGTGGAGCCGAGCTCCAGAGA 540
QY 1110 AGGAGGAGAACCAAGAAAGATGATATATATATCTGCTTTGGATCGAAGAGCGGTATC 1169
Db 541 AGGAGGAGAACCAAGAAAGATGATATATATATCTGCTTTGGATCGAAGAGCGGTATC 600
QY 1170 CCACTGTGAGGAGCAGGAGCCGCTGCTCCGGAATGATGATGAGCCGAGCCGAGCT 1228
Db 601 CCACTGTGAGGAGCAGGAGCCGCTGCTCCGGAATGATGATGAGCCGAGCCGAGCT 660
QY 1229 GTGATATCTCCATGCTGAGAGCCGCTCAGAGAGAGCCGAGAGAGTGCATGAGA 1288
Db 661 GTGATATCTCCATGCTGAGAGCCGCTCAGAGAGAGCCGAGAGAGTGCATGAGA 720
QY 1289 GTCTGTGAGCAGTCCGAGTCCGAGGAGGAGTCCCTGATCCAGCCGAGCGGCTTG 1348
Db 721 GTCTGTGAGCAGTCCGAGTCCGAGGAGGAGTCCCTGATCCAGCCGAGCGGCTTG 780
QY 1349 GAGATGAGCCAG 1360
Db 781 GCCTTGCCAGAG 792

RESULT 3
LOCUS BF529743
DEFINITION 831 bp mRNA linear EST 11-DEC-2000
ACCESSION BF529743
VERSION BF529743
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 831)
NIH-MGC http://mgi.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs@mail.nih.gov
Tissue Procurement: David N. Louis, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Inyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNI at:
http://image.llnl.gov
Plate: LLAM9495 row: c column: 01
High quality sequence stop: 736.
Location/Qualifiers
1. 831
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:418168"
/clone_1db="NCI-CCAP_Brn67"
/tissue_type="anaplastic oligodendroglioma with 1p/19q

loss"
/lab_host="DH10B (71 phage-resistant)"
/note="Organ: brain; Vector: pCMV-SpRTE; Site: 1: Not;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCI-CCAP library."
BASE COUNT 185 a 237 c 258 g 151 t
ORIGIN
Query Match 21.9%; Score 651.2; DB 12; Length 831;
Best Local Similarity 96.3%; Pred. No. 5e-80;
Matches 709; Conservative 0; Mismatches 23; Indels 4; Gaps 4;
QY 403 GTACTGTTCTGTGAGCAGCTCTGAGGAGGAGTGAATTCATCTGATTAAGAGAGG 462
Db 44 GTACTGTTCTGTGAGCAGCTCTGAGGAGGAGTGAATTCATCTGATTAAGAGAGG 103
QY 463 GAGACTGAGCCCAA -GGAGGCCGAAAGTTCTCCGCAATTTGTCTGCTGACT 521
Db 104 GAGACTGAGCCCAAAGGAGGCCGAAAGTTCTCCGCAAG -TGTGTCTGCTGACT 162
QY 522 TGTGCCAGCTACTCATCTGTCAGAGAGACTTAAGCCCGAGAACCTGCTTTGGATG 581
Db 163 TGTGCCAGCTACTCATCTGTCAGAGAGACTTAAGCCCGAGAACCTGCTTTGGATG 222
QY 582 AGAAAAACAATCCGATTCAGACTTCGAGATGCGCTCCCTGCGAGGTGGGAGAGCC 641
Db 223 AGAAAAACAATCCGATTCAGACTTCGAGATGCGGCTCCCTGCGAGGTGGGAGAGCC 282
QY 642 TCTGTGAGACAGTCTGCGGCTCCCTCATTTGCTGTCCAGAGGTGATTAAGGGGAAA 701
Db 283 TCTGTGAGACAGTCTGCGGCTCCCTCATTTGCTGTCCAGAGGTGATTAAGGGGAAA 342
QY 702 AATATGAGCGCGCGGAGAGATGAGAGTGTGAGTATCTCTGCGCCCTGCTG 761
Db 343 AATATGAGCGCGCGGAGAGATGAGAGTGTGAGTATCTCTGCGCCCTGCTG 402
QY 762 TGGGGGCTTGCCCTTTGATGAGAGACAATCCGCGAGCTCTGAGAGAGTGAAGCGG 821
Db 403 TGGGGGCTTGCCCTTTGATGAGAGACAATCCGCGAGCTCTGAGAGAGTGAAGCGG 462
QY 822 GCGTCTTCACATGCCCCCAGTCTTCTCCAGATGTCAGAGAGCTCTGAGAGAGAGA 881
Db 463 GCGTCTTCACATGCCCCCAGTCTTCTCCAGATGTCAGAGAGCTCTGAGAGAGAGA 522
QY 882 TCGAATGAGAGCCCGAAGAAAGGCTCAGTGTGAGCAATTCAGAAATCATCTTGATCC 941
Db 523 TCGAATGAGAGCCCGAAGAAAGGCTCAGTGTGAGCAATTCAGAAATCATCTTGATCC 582
QY 942 TAGGCGGGAACACAGAGCCAGACCCGCTGAGAGCCAGCCCTGCGCGGCTAGCCA 1001
Db 583 TAGGCGGGAACACAGAGCCAGACCCGCTGAGAGCCAGCCCTGCGCGGCTAGCCA 642
QY 1002 TGGCG -AGCTGCTCATTCACAGAGAGTGTGAGAGCCGAGCTCTGAGAGAGTGCATCA 1060
Db 643 TGGCGAAGCTGCTCATTCACAGAGAGTGTGAGAGCCGAGCTCTGAGAGAGTGCATCA 701
QY 1061 CTGGGCTGCTTCAGAGAGCCGAGAGAGTGTGAGAGCCGAGCTCTGAGAGAGTGCATCA 1120
Db 702 CTGGGCTGCTTCAGAGAGCCGAGAGAGTGTGAGAGCCGAGCTCTGAGAGAGTGCATCA 761
QY 1121 CAAGAAAGATGATAT 1136
Db 762 ACAAAATGTTTATAT 777

RESULT 4
LOCUS BM675364/c
DEFINITION 678 bp mRNA linear EST 27-FEB-2002
ACCESSION BM675364
VERSION BM675364.1 GI:18985262
UI-E-E10-ahr-p-12-0-UI-s1 UI-E-E10-ahr-p-12-0-UI-3', mRNA sequence.

EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 678)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hagaman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA library arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 201-243, >(TAA)n#Simple_repeat
Seq primer: M13 Forward
POLYA-yes.

/clone="IMAGE:4932680"
 /tissue="NCI_GAP_Brn67"
 /tissue_type="anaplastic_oligodendroglioma_with_1p/19q_loss"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: brain; Vector: PCMV-SPORT6; Site: 1; Motif: Site: 2; Salt: Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.3 kb. Constructed by Life Technologies. Note: this is a NCI_GAP library."
 BASE COUNT 145 a 194 c 223 g 96 t
 ORIGIN

Query Match 20.2%; Score 602.2; DB 12; Length 658;
 Best Local Similarity 98.4%; Pred. No. 2.9e-73;

Matches 619; Conservative 0; Mismatches 8; Indels 2; Gaps 1;

QY 944 GGGGGAACACGAGCAGACCCGCTGCTGAGCAGCCCGCTGCGCGGCTAGCCATG 1003
 DB 1 GGGGGGAACACGAGCAGACCCGCTGCTGAGCAGCCCGCTGCGCGGCTAGCCATG 60
 QY 1004 CGGAGCCTGCTCCTCAACGAGAGCTGACCCGAGCTCTAGAGAGATGCTACTG 1063
 DB 61 CGGAGCCTGCTCCTCAACGAGAGCTGACCCGAGCTCTAGAGAGATGCTACTG 120
 QY 1064 GCGTCTTCAAGGAGCCGAGAGAGCTGCTGAGCTGCGCAGTGAAGAGAGACCA 1123
 DB 121 GCGTCTTCAAGGAGCCGAGAGAGCTGCTGAGCTGCGCAGTGAAGAGAGACCA 180
 QY 1124 GAAAGATGATATATATATGCTTTGATGGAAGAGAGCTGATCCAGCTGAGAGC 1183
 DB 181 GAAAGATGATATATATATGCTTTGATGGAAGAGAGCTGATCCAGCTGAGAGC 240
 QY 1184 CAGAGCCTGCTCCTCCTGGAATGATGTTGACCCCGGAAAGCTGTGATTCCTCATG 1243
 DB 241 CAGAGCCTGCTCCTCCTGGAATGATGTTGACCCCGGAAAGCTGTGATTCCTCATG 300
 QY 1244 CTGAGCCTGCTCCTCCTGGAATGATGTTGACCCCGGAAAGCTGTGATTCCTCATG 1303
 DB 301 CTGAGCCTGCTCCTCCTGGAATGATGTTGACCCCGGAAAGCTGTGATTCCTCATG 360
 QY 1304 GATGCGGAGGCTGCTCCTGCTGACCCGAGAGGCTGCTGAGATGAGCCAGCAGC 1363
 DB 361 GATGCGGAGGCTGCTCCTGCTGACCCGAGAGGCTGCTGAGATGAGCCAGCAGC 420
 QY 1364 AGCCAGAGATCCGCTGATGAGAGCTGACAGGCTGCTGCTGACCCCTGTAAGC 1423
 DB 421 AGCCAGAGATCCGCTGATGAGAGCTGACAGGCTGCTGCTGACCCCTGTAAGC 480
 QY 1424 AGCCAGAGATCCGCTGATGAGAGCTGACAGGCTGCTGCTGACCCCTGTAAGC 1483
 DB 481 AGCCAGAGATCCGCTGATGAGAGCTGACAGGCTGCTGCTGACCCCTGTAAGC 540
 QY 1484 GGGGGGAGCTCCCGATTCCTTTTCAACGAGAGCTGCTGCTGAGGAGGAGGCTGAG 1543
 DB 541 GGGGGGAGCTCCCGATTCCTTTTCAACGAGAGCTGCTGCTGAGGAGGAGGCTGAG 598
 QY 1544 GGGCGGGGAGAGAGCCCGCCCGCCAG 1572
 DB 599 GGGCGGGGAGAGAGCCCGCCCGCCAG 627

RESULT 6
 BM948524 698 bp mRNA linear EST 14-MAR-2002
 LOCUS UI-M-EGOP-lve-1-10-0-UI.r1.NH.BMAP_EGop Mus musculus cdna clone
 DEFINITION IMAGE:5691033 5', mRNA sequence.
 ACCESSION BM948524
 VERSION BM948524.1 GI:19432115
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathia; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 698)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES
 source
 location/Qualifiers
 1..698
 /organism="Mus musculus"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5691033"
 /clone_1ib="NIH_BMAP_EGop"
 /tissue_type="whole brain"
 /dev_stage="embryo 18.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /note="Organ: brain; Vector: pYX-Asc; Site: 1; Ecov I; Site: 2; Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is CAGCCAGCAGC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemh Clin, Ph.D., program coordinator."

BASE COUNT 165 a 190 c 209 g 134 t
 ORIGIN

Query Match 19.9%; Score 592.4; DB 14; Length 698;
 Best Local Similarity 90.5%; Pred. No. 6.1e-72;
 Matches 632; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 417 AGCAGCTCTGGGGGTGAGCTATGCACTACTGTTAAGAGAGGAGACTGACGCCA 476
 DB 1 AGCAGCTCTGGGGGTGAGCTATGCACTACTGTTAAGAGAGGAGACTGACGCCA 60
 QY 477 AGGAGGCGCGAAGTTCTTCCGCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 536
 DB 61 AGGAGGCGCGAAGTTCTTCCGCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 QY 537 CCATCTGCCACAGAGACTAAAGCCGAGAACTGCTTTTGGATGAGAAAAACAATCC 596
 DB 121 CCATCTGCCACAGAGACTTGAAGCCAGAGAACTGCTTTTGGATGAGAAAAACAATCC 180
 QY 597 GCATTCGAGACTTCGCGATGCGCTCCCTGCAAGTGGGGGAGACGCTCTGAGACCACT 656
 DB 181 GCATTCGAGACTTCGCGATGCGCTCCCTGCAAGTGGGGGAGACGCTCTGAGACCACT 240
 QY 657 GCGGTCGCCCGCAATTAATGCGTCCAGAGGATTAAGGGGAGAAAAATGATGGCGCCG 716
 DB 241 GCGGTCGCCCGCAATTAATGCGTCCAGAGGATTAAGGGGAGAAAAATGATGGCGCCG 300
 QY 717 GGGCAGACATGTGAGAGCTGTGAGATCACTCTTCCCTCTGCTGAGGCTGCTGCTGCT 776
 DB 301 GGGCAGACATGTGAGAGCTGTGAGATCACTCTTCCCTCTGCTGAGGCTGCTGCTGCT 360
 QY 777 TTGATGACGCAACCTCCGCGAGCTGCTGAGAAAGGTGAAGCGGGCGCTCTTCCACATGC 836

```

Db 361 TCGATACGCAACCTGCGCCAGCTACTGAGAGGTAAGCGGGGTCTTCCACATGCG 420
QY 837 CCCACTTCACTCTCCAGATGTCAGAGCCCTCTGAGGGGAGATGATGAGAGCCG 896
Db 421 CTCACCTCACTCTCCAGATGTCAGAGCCCTCTGAGGGGATGATGATGAGAGCCG 480
QY 897 AAAAAAGGCTCACTCTGAGAGCAATTCAGAAACATCTTGTGCTACCTGAGCCGGAACACG 956
Db 481 AGAAAAAGGCTCACTCTGAGAGCAATTCAGAAACATCTTGTGCTACCTGAGCCGGAACACG 540
QY 957 AGCCAGACCCGCTGCTGAGAGCCAGCCCTGCGCCGCGGTAGCATGCGGAGCCGCTG 1016
Db 541 AACACAGACCTTCTGCTGAGAGCCAGCCGCGCGGTAGCATGCGGAGCCGCTG 600
QY 1017 CCACAGAGAGCTGAGAGCCGCGAGCTCTAGAGAGCATGAGTCTGAGGCTGCTTCAAGG 1076
Db 601 CCATGTCGAGCTGAGAGCCGCTGAGTCTGAGAGAGCATGAGGCTGCTGCTTCAAGG 660
QY 1077 ACCGCGAGAGCTGAGTCTGAGAGCTGAGGCTGAGAGG 1114
Db 661 ACCGCGAGGCTGAGAGAGAGTCTGAGAGCGAGAG 698

```

```

RESULT 7
BM944329 742 bp mRNA linear EST 14-MAR-2002
LOCUS BM944329
DEFINITION UI-M-EH0P-bvq-d-01-0-UI.r1 NIH_BMAP_EH0P Mus musculus cDNA clone
IMAGE:5695440 5', mRNA sequence.
ACCESSION BM944329
VERSION BM944329.1 GI:19427914
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
1 (bases 1 to 742)
NIH-MGC http://mgs.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

```

```

FEATURES
Source
Seq primer: pYX-5.
Location/Qualifiers
1..742
/organism="Mus musculus"
/strain="C57Bl/6"
/db_xref="taxon:10090"
/clone="IMAGE:5695440"
/clone_1ib="NIH_BMAP_EH0P"
/tissue_type="whole brain"
/dev_stage="embryo 18.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/note="Organ: brain; Vector: pYX-Asc; Site: 1: EcoR I;
Site: 2: Not I; The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
, is CAGCCAGCAGC. This library was created for the

```

University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institute of Mental Health
(NIMH), Hemlin Chin, Ph.D., program coordinator."

BASE COUNT 179 a 197 c 219 g 145 t 2 others

Query Match 19.3%; Score 574.8; DB 14; Length 742;
Best Local Similarity 89.4%; Pred. No. 1.5e-69;
Matches 663; Conservative 0; Mismatches 74; Indels 5; Gaps 4;

```

QY 527 CACAGTACTCATCTGCGACAGAGACCTTAACCCGAGAACCTGCTTTGGATGAGAAA 586
Db 2 CATAGTACTCATCTGCGACAGAGACCTTAACCCGAGAACCTGCTTTGGATGAGAAA 61
QY 587 AACACATCCGATTCAGACCTTGGCATGCTGCTGCTGAGAGTGGGGAGAGCCCTCTG 646
Db 62 AACACATCCGATTCAGACCTTGGCATGCTGCTGAGAGTGGGGAGAGCCCTCTG 121
QY 647 GAGACCAGCTGCGGGTCCGCCCTTATGCGTGTCCAGAGGTGATTAAGGGGAAAAATAT 706
Db 122 GAGACCAGCTGCGGGTCCGCCCTTATGCGTGTCCAGAGGTGATTAAGGGGAAAAATAT 181
QY 707 GATGGCGCGCGGGGAGACATGTGGAGCTGTGGAGTATCTCTTCCGCTGCTGAGGG 766
Db 182 GATGGCGCGCGGGGAGACATGTGGAGCTGTGGAGTATCTCTTCCGCTGCTGAGGG 241
QY 767 GCTCTGCCCTTTGATGACACAACTCCGCGACCTGCTGAGAGAGGAAAGGGGGCTG 826
Db 242 GCACGTGCTTGTGATGACACAACTCCGCGACCTGCTGAGAGAGGAAAGGGGGCTG 301
QY 827 TTCCACATGCCCTCACTTCTCTCCAGATGTCAGAGCCCTCTCTGAGAGGATGATCGAA 886
Db 302 TTCCACATGCCCTCACTTCTCTCCAGATGTCAGAGCCCTCTCTGAGAGGATGATGAA 361
QY 887 GTGGAGCCCGGAAAAAGGCTGAGTGTGGAGCAATTCAGAAACATCTTGGTA--CCTAG 944
Db 362 GTGGAGCCCGGAAAAAGGCTGAGTGTGGAGCAATTCAGAAACATCTTGGTA--CCTAG 421
QY 945 GCGGGAACACGACGAGACCCGCTGCTGAGAGCCACCCCTGCGCGCGGAGAGCATGCG 1004
Db 422 GCGGGAACACGACGAGACCCGCTGCTGAGAGCCACCCCTGCGCGCGGAGAGCATGCG 481
QY 1005 GAGCCCTGCCATTCAGAGAGAGAGTGGAGCCCGGAGCTCTAGAGAGCATGCACTGG 1064
Db 482 GTAGCGCTGCTTCTGATGAGAGTGGAGCCCTGAGTGTGGAAGACATGGCGTCTGTGG 541
QY 1065 GCTGCTTCAAGGAGCGGAGAGGCTGATCGAGTGGCGGAGTGGAGAGAGAACCAAG 1124
Db 542 GCTGCTTCAAGAGACCGCGAGCGGCTGACAGAGAACTGCGAGAGAGAGAACCAAG 601
QY 1125 AAAAGATGATATATATATCTGC--TTTGGATCGAGAGAGCGGTATCCAGCTGAGAGAC 1183
Db 602 AAAAGATGATATATATATATGCTTTTGGATCGAGAGAGCGGTATCTAGCTGAGAGAC 661
QY 1184 CAGGACCTGCTCCCGCGGATGATGTTGACCC--CCCCGGAAGCGTGTGATTCCTCAT 1242
Db 662 CAGGAC--CTGCTCTCTCTCGGATGATGTTGACCCACCTCGTAGCGTGTGATTCCTCAT 720
QY 1243 GCTGAGCGGTGACGAGGAGCGG 1264
Db 721 GCTGAGCANACAGGGAACCG 742

```

```

RESULT 8
B0178547 671 bp mRNA linear EST 30-APR-2002
LOCUS B0178547
DEFINITION UI-M-EV0-bws-b-03-0-UI.r1 NIH_BMAP_EV0 Mus musculus cDNA clone
IMAGE:5701154 5', mRNA sequence.
ACCESSION B0178547
VERSION B0178547.1 GI:20354039
KEYWORDS EST.
SOURCE house mouse.

```


cloned directionally into pTX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TCAGAGAGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator.

BASE COUNT 153 a 185 c 195 g 125 t
ORIGIN

Query Match 18.1%; Score 540.4; DB 14; Length 658;
Best Local Similarity 89.8%; Pred. No. 8e-65;
Matches 591; Conservative 0; Mismatches 66; Indels 1; Gaps 1;

QY 820 GGGGCTTCCACATGCCCCCCTTCATTCAGATGCCAGAGCCCTCCAGGGGAT 879
DB 2 GGGGCTTCCACATGCCCCCCTTCATTCAGATGCCAGAGCCCTCCAGGGGAT 61
QY 880 GATCGAATGAGCCGCAAAAGGCTGAGTGGAGCAATTCGAAACATCTTGGTA 939
DB 62 GATTGAAGTGGAGCCGAGAAAGG-TCAGTCTGAGCAAAATTCAGAAACATCTTGGTA 120
QY 940 CTTAGCGGGGAAACAGACAGAGCCGCTGGAGCCAGCCCTGGCGCGGGTAC 999
DB 121 TCTGGGGGGAACACAGACAGACCCCTGCTGGAGCCAGCCCGCAGAGTAC 180
QY 1000 CATCGGAGCCCTCCATCCAAAGAGAGCTGAGCCCGAGCTCTAGAGCATGGATC 1059
DB 181 CATGCTTACGCTGCTTCATTCAGATGCCAGAGCCCTGAGCTTGAAGAGCGCTC 240
QY 1060 ACTGGGCTGCTTCCAGAGAGCCGAGAGCTGATCCGAGCTGGCCATGAGAGAGAA 1119
DB 241 TCTGGGCTGCTTCCAGAGAGCCGAGAGCTGATCCGAGAGCTGGCCATGAGAGAGAA 300
QY 1120 CCAAGAAAGATGATATATATATCTGCTTTGGATCGAAGAGCGATATCCAGCTGTA 1179
DB 301 CCAAGAAAGATGATATATATATCTGCTTTGGATCGAAGAGCGATATCCAGCTGTA 360
QY 1180 GGACAGAGAGCTGCTCCCGGAAATGATGTTGACCCCGCGGAAAGCTGTGATTC 1239
DB 361 AGACAGAGAGCTGCTCCCGGAAATGATGTTGACCCCGCGGAAAGCTGTGATTC 420
QY 1240 CATGCTGAGCTGAGAGAGAGCGGAGAGCGGAGAGCTGAGAGCTGAGAT 1299
DB 421 CATGCTGAGCTGAGAGAGAGCGGAGAGCGGAGAGCTGAGAGCTGAGAT 480
QY 1300 CACGAGTCCCGGGGAGTGGCTCCCTGTACCCAGCCGAGCGGCTTGAAGATGCCCA 1359
DB 481 CACGAGTCCCGGGGAGTGGCTCCCTGTACCCAGCCGAGCGGCTTGAAGATGCCCA 540
QY 1360 GCACAGCAGAGATCCCGTACGCTGAGAGCTCCAGCGGCTGTCTCCAGCCCTCT 1419
DB 541 GCACAGCAGAGATCCCGTACGCTGAGAGCTCCAGCGGCTGTCTCCAGCCCTCT 600
QY 1420 AAGCAGCCCAAGAGAGTCCGCTTCTTCAACCGGAGCGGGGGTGGAGATGAG 1477
DB 601 GAGCAGCCCAAGAGAGTCCGCTTCTTCTTCAACCGGAGCGGCGTGGAGATGAG 658

RESULT 10
AL538014 836 bp mRNA linear EST 13-FEB-2001
LOCUS AL538014 LTI_FL013_FBrn1 Homo sapiens cDNA clone CS0DF029YC21 5
DEFINITION prime, mRNA sequence.
ACCESSION AL538014
VERSION AL538014.1 GI:12801507
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 836)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polyes, D.

TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr.
Location/Qualifiers

FEATURES

source

1. 836

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0DF029YC21"

/dev_stage="LTI_FL013_FBrn1"

/dev_stage="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)"

/lab_host="DH10B"

/note="Organ: Retal brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@life.com URL : http://fulllength.invitrogen.com"

BASE COUNT 181 a 230 c 268 g 152 t 5 others
ORIGIN

Query Match 17.7%; Score 527.6; DB 9; Length 836;
Best Local Similarity 81.8%; Pred. No. 3.e-63;
Matches 630; Conservative 4; Mismatches 138; Indels 2; Gaps 2;

QY 164 CACGCCCAATATGTCGGCCCTTATCGGCTGAGAGAGCGTGGGCAAGACAGACAGG 223
DB 65 CACGCCCAATATGTCGGCCCTTATCGGCTGAGAGAGCGTGGGCAAGACAGAGT 124
QY 224 CTGCTTAATCTGGGCTTCACTGATCAGAGGTCAGAGTCCCATCAAGATGTAAC 283
DB 125 CTGCTTAATCTGGGCTTCACTGATCAGAGGTCAGAGTCCCATCAAGATGTAAC 184
QY 284 CGGAGAGCTGTCGGAGTGGGCTGATGAAGTGAAGCGGAGATCCCATCTGAG 343
DB 185 CGGAGAGCTGTCGGAGTGGGCTGATGAAGTGAAGCGGAGATCCCATCTGAG 244
QY 344 CTGATGAAACCCACATGCTCTCAAGCTCCAGAGCTCTCAAGAAAGAAATATTTG 403
DB 245 CTGATGAAACCCACATGCTCTCAAGCTCCAGAGCTCTCAAGAAAGAAATATTTG 304
QY 404 TACCTGTTTGAACAGAGCTCTGGGGGTGAGCTATTGCACTACCTGTAAAGAGGG 463
DB 305 TACCTGTTTGAACAGAGCTCTGGGGGTGAGCTATTGCACTACCTGTAAAGAGGG 364
QY 464 AGACTAGCGCCCAAGGAGCGCCGAAGTCTCCGCGAGATGTCGGCTGAGCTTC 523
DB 365 AGACTAGCGCCCAAGGAGCGCCGAAGTCTCCGCGAGATGTCGGCTGAGCTTC 424
QY 524 TGGCAGAGTACTCATCTGACAGAGACCTAAAGCCGAGAGCTCTTTGGATGAG 583
DB 425 TGGCAGAGTACTCATCTGACAGAGACCTAAAGCCGAGAGCTCTTTGGATGAG 484
QY 584 AAAAACAATCCGCAATTCGACACTTGGCATGGCTCCCTGCAAGTGGGGGACAGCTC 643
DB 485 AAGAACAAATCCGCAATTCGACACTTGGCATGGCTCCCTGCAAGTGGGGGACAGCTC 544
QY 644 CTGAGACAGAGTGGGGTCCCAATTAAGTGGTCCAGAGGATTAAGGGGAGAAA 703
DB 545 TTGAGACAGAGTGGGGTCCCAATTAAGTGGTCCAGAGGATTAAGGGGAGAAA 604
QY 704 TATGATGCGCGCGGAGCATGAGAGCTGTGAGTCACTCTTCCGCTGCTGCTG 763
DB 605 TATGATGCGCGGAGCATGAGAGCTGTGAGTCACTCTTCCGCTGCTGCTGCTG 663
QY 764 GGGGCTTCCCTTTGATGACGAAACCTTCGCGAGCTGTGAGAGAGTGAACGGGGC 823

OY	824	GTCCTCCAGATGCCCCCACTTACATCCTCCAGATTGCCAGAGCGCTCTGGAGGAATGATGC	883
DJ	664	GGGGCTGTGCCCCTTGAGAGATGATCAACAATTTGGCACAA-CGTCTGGAGAAGGTGAAGCGGGGC	722
OY	824		
Dj	723		
OY	884	GAGTGGAGCCCCGAAAAAAGGCTCAGTCTGGAGCAAATTCAGAAAATCCTTGG	937
Dd	783		
		GAGGTGAGCGCCGCGACGGCGCTCAGCGTAGAGCATTTAGAAAACACATAATKG	836
		:	
RESULT	11	BFS29630	936 bp mRNA linear EST 11-DEC-2000
LOCUS	BFS29630	60204385F01 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4181581	
DEFINITION		5', mRNA sequence.	
ACCESSION		BFS29630	
VERSION		BFS29630.1 GI:11616993	
KEYWORDS		EST.	
SOURCE		human.	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE		1 (bases 1 to 936)	
JOURNAL		NIH-MGC http://mgc.nci.nih.gov/.	
COMMENT		National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabs-retail.nih.gov Tissue Procurement: David N. Louis, M.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNI) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNI at: http://image.llnl.gov Plate: LLNMG494 row: n column: 14 High quality sequence start: 8 High quality sequence stop: 568. Location/Qualifiers 1..936 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4181581" /clone_1id="NCI_CGAP_Brn67" /tissue_type="anaplastic oligodendroglioma with lp/1p/q loss" /lab_host="DH10B (TI phage-resistant)" /note="Organ: brain; Vector: pCMV-SPOUT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.3 kb. Constructed by Life Technologies. Note: This is a NCI_CGAP library."	
BASE COUNT		286 a 238 c 318 g 94 t	
ORIGIN			
Query Match		17.5%; Score 522.4; DB 12; Length 936;	
Best Local Similarity		.97.68; Pred. No. 1.8e-62;	
Matches 562; Conservative 0; Mismatches 11; Indels 3; Gaps 3			
OY	986	GGCCCGCGGTAGACCATCGAGACTTCGCATCCACGAGAGCTGAGCCCGACGTCTTA	1045
Dd	5	GGCCGCGGTAGACCATCGAGACTTCGCATCCACGAGAGCTGAGCCCGACGTCTTA	64
OY	1046	GAGAGCATGGCATCTCTGGCTGCTTCAAGGAGCCGAGAGGCTGCATCGCAGCTCGCC	1105
Dd	65	GAGAGCATGGCATCTCTGGCTGCTTCAAGGAGCCGAGAGGCTGCATCGCAGCTCGCC	124
OY	1106	AGTGAGAGGAGAACCAAGAAAAGATGATATTTATCTGCTTTTGATGGAAGAGCGG	1165
Dd	125	AGTGAGGAGGAGAACCAAGAAAAGATGATATTTATCTGCT-TTTGATGGAAGAGCGG	183
OY	1166	TATCCACAGCTGTGAGGACCAAGACCTGCTCCCGGAATGATGTTTGACCCCGCGGAAG	1225

Db	184	TATCCACACCTGTGAGAGACCAAGACCTGCTCCCTCCCGAATGATGTGACCCCCCGGAAG	243
QY	1226	CCTGGATTTCTCCCANCTAGCCCTACCGGAAAGCGGCGACCAAGACGGAATGTCATG	1285
Db	244	CGTGGATTTCTCCCATCTAGCCCTACGGAAGCGGCGA-CAGACGGGAATGTCATG	302
QY	1286	GAAGTCCAGACATCACCGATCCCGGGGGTGGTGGCTCCCTGTACCAACCGAGCGCC	1345
Db	303	GAAGTCCAGACATCACCGATCCCGGGGGTGGTGGCTCCCTGTACCAACCGAGCGCC	362
QY	1346	TTTGAGATGGCCCGACACAGACAGATCCGTAGCTGATGAGCCTCCACGSGTCTG	1405
Db	363	TTTGAGATGGCCCGACACAG-CAGAGATCCGTAGCTGATGAGCCTCCACGSGTCTG	421
QY	1406	TCCTCCACCTCTTAGAGAGCCCAAGAGTCCGGCTTTCTCTTTTACCGGAGCGGGG	1465
Db	422	TCCTCCACCTCTTAGAGAGCCCAAGAGTCCGGCTTTCTCTTTTACCGGAGCGGGG	481
QY	1466	GCTGAGATGAGGCTCGAGGGGGGCTCCCGACTTCCAAAACGACAGACGCTGCTTCT	1525
Db	482	GCTGAGATGAGGCTCGAGGGGGGCTCCCGACTTCCAAAACGACAGACGCTGCTTCT	541
QY	1526	CGGGGCCCCAGGGGTGGGGGGCGCGGGGAGACAGCCC	1561
Db	542	TCGGGGCCCCAGGGGTGGGGGGCGCGGAAACAGCAC	577

RESULT 12	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
BO086295	510 bp	mRNA	linear	EST 29-APR-2002		human.	Homo sapiens					
						human.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
							Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
							1 (bases 1 to 510)					
							Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Keastner,K.,					
							Leimiska,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S.,					
							Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,D., Bistain,A.,					
							Schmitt,A., Theisling,B., Ritter,E., Ronko,I., Bennett,J., Cardenas					
							,M., Gibbons,M., McCann,R., Cole,R., Tsagarashvili,R., Williams,T.,					
							Jackson,Y. and Bowers,Y.					
							Endocrine Pancreas Consortium					
							Unpublished (2000)					
							Contact: Douglas Melton, Klaus H. Keastner, & Hiroshi Inoue					
							Endocrine Pancreas Consortium					
							Harvard University, Howard Hughes Medical Institute					
							Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,					
							MA 02138					
							Tel: 617-495-1812					
							Fax: 617-495-8557					
							Email: dmelton@biohp.harvard.edu					
							Library was constructed by Dr. Douglas Melton DNA sequencing by:					
							Washington University Genome Sequencing Center this clone is					
							available royalty-free through INM, please contact the IMAGE					
							consortium (info@image.llnl.gov) for further information					
							Seq primer: -40RP from Gibco					
							High quality sequence stop: 445.					
							Location/Qualifiers					
							1. 510					
							/organism="Homo sapiens"					
							/db_xref="taxon:9606"					
							/clone="IMAGE:6135408"					
							/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"					
							/sex="Both"					
							/tissue_type="Islets of Langerhans"					
							/dev_stage="Adult"					

FEATURES	Source
----------	--------

```

/lab.host="DH10B"
/Note="Organ: Pancreas; Vector: pSPORT1; Site: 1; Not 1;
Site: 2; Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an EcoT of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

BASE COUNT      110 a      171 c      136 g      93 t
ORIGIN

Query Match      17.1%; Score 510; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.4e-60;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1777 TCGCTTCAACGGCGGAGATGAGAGTCCCTACCGGTGAGAGATGCCAGCTTGACGCC 1836
DB 1 TCGCTTCAACGGCGGAGATGAGAGTCCCTACCGGTGAGAGATGCCAGCTTGACGCC 60
QY 1837 AGAGTCTCCCGGAGAGTGGCAAAACGCTCTGTTGCGGAACTTCATCTCTTGACAA 1896
DB 61 AGAGTCTCCCGGAGAGTGGCAAAACGCTCTGTTGCGGAACTTCATCTCTTGACAA 120
QY 1897 AGAAGAACAAATTTCTCTGCTGCTAAAGAACAACTCTGACAGCATCAAGACAGCAT 1956
DB 121 AGAAGAACAAATTTCTCTGCTGCTAAAGAACAACTCTGACAGCATCAAGACAGCAT 180
QY 1957 CGTCACATGCTTCTGTCATCCCGACGCTGAGTCACAGTGTGTCACAGACAGCTT 2016
DB 181 CGTCACATGCTTCTGTCATCCCGACGCTGAGTCACAGTGTGTCACAGACAGCTT 240
QY 2017 CAGGCGCGAGTACAGAGCCAGTGGCGGCCCTCCCTCTTCCAAAAGCCCTGCTTCCA 2076
DB 241 CAGGCGCGAGTACAGAGCCAGTGGCGGCCCTCCCTCTTCCAAAAGCCCTGCTTCCA 300
QY 2077 GGTGGACATCAGTCTCTGAGGTCAGAGCCCTCCCGGACGAGGAGGAGGAGG 2136
DB 301 GGTGGACATCAGTCTCTGAGGTCAGAGCCCTCCCGGACGAGGAGGAGGAGGAGG 360
QY 2137 TGGTGACATCTACTCCGTCACCTTCACTCATCTGAGTGGGTCCCGGCTGCTCAAGCG 2196
DB 361 TGGTGACATCTACTCCGTCACCTTCACTCATCTGAGTGGGTCCCGGCTGCTCAAGCG 420
QY 2197 AGTGTGAGAGACCATCAGGACAGCTCTGAGACATCATGACGCGCTCCGTCAGAGC 2256
DB 421 AGTGTGAGAGACCATCAGGACAGCTCTGAGACATCATGACGCGCTCCGTCAGAGC 480
QY 2257 CCTGGCAGACGAGAAAGAGGGGGCCGAC 2286
DB 481 CCTGGCAGACGAGAAAGAGGGGGCCGAC 510

RESULT 13
LOCUS      BF529174      898 bp      mRNA      linear      EST 11-DEC-2000
DEFINITION 602042051r1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4179828
5', mRNA sequence.
ACCESSION  BF529174
VERSION     BF529174.1 GI:11616455
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 898)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/

```

```

TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
CONTACT    Robert Strausberg, Ph.D.
COMMENT    Email: cgaphs-remail.nih.gov
            Tissue Procurement: David N. Louis, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LIML at:
            http://image.llnl.gov
            Plate: LHM9490 row: e column: 13
            High quality sequence stop: 587.

FEATURES
    source
        1..898
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:4179828"
        /clone_1ib="NCI_CGAP_Brn67"
        /tissue_type="anaplastic oligodendroglioma with 1p/19q
        loss"
        /lab.host="DH10B (T1 phage-resistant)"
        /note="Organ: brain; Vector: pCMV-SPORT6; Site: 1; Not 1;
        Site: 2; Sal1; Cloned unidirectionally. Primer: Oligo dT.
        Average insert size 2.3 kb. Constructed by Life
        Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT      154 a      275 c      280 g      189 t
ORIGIN

Query Match      16.5%; Score 490.8; DB 12; Length 898;
Best Local Similarity 98.1%; Pred. No. 3.9e-58;
Matches 518; Conservative 0; Mismatches 7; Indels 3; Gaps 2;

QY 403 GTACCTGTTTGGAGACAGTCTCGGGGGGTGAGCTTATTCAGTACTGGTAAAGAGG 462
DB 89 GTACCTGTTTGGAGACAGTCTCGGGGGGTGAGCTTATTCAGTACTGGTAAAGAGG 148
QY 463 GAGACTGAGCCCAAGAGAGGCGGAAAGTTCTTCCGACAGATTGTCTCGTGGACTT 522
DB 149 GAGACTGAGCCCAAGAGAGGCGGAAAGTTCTTCCGACAGATTGTCTCGTGGACTT 208
QY 523 CTGCCACACTTCTCATCTGCCACAGACGCTTAAAGCCCGAGAACCTGTTTGATGA 582
DB 209 CTGCCACACTTCTCATCTGCCACAGACGCTTAAAGCCCGAGAACCTGTTTGATGA 268
QY 583 GAAAAACAACATCCGATTTGACAGATTTGGCATGGGTCCCTGAGTGGGGACACCT 642
DB 269 GAAAAACAACATCCGATTTGACAGATTTGGCATGGGTCCCTGAGTGGGGACACCT 328
QY 643 CTGGAGACCACTGGGGTCCCGCCATTATGCGTTCAGAGGTATTAAAGGGGAAAA 702
DB 329 CTGGAGACCACTGGGGTCCCGCCATTATGCGTTCAGAGGTATTAAAGGGGAAAA 388
QY 703 ATATGATGGCCCGCGGAGACATGTGAGCTGTGAGATCATCTCTTCCCTGCTGCT 762
DB 389 ATATGATGGCCCGCGGAGACATGTGAGCTGTGAGATCATCTCTTCCCTGCTGCT 448
QY 763 GGGGGCTGCGCTTGTAGTACAGACAACCTCCGCGAGCTGTGAGAGAGTGAAGGGG 822
DB 449 GGGGGCTGCGCTTGTAGTACAGACAACCTCCGCGAGCTGTGAGAGAGTGAAGGGG 508
QY 823 CGTCTTCACATGCCCACTTCAATTCCTCCAGATTGCCAGAGGCTCTGAGGGGAAT -G 880
DB 509 CGTCTTCACATGCCCACTTCAATTCCTCCAGATTGCCAGAGGCTCTGAGGGGAATGA 568
QY 881 ATCGAAGTGGAGCCCG -AAAAAAGGCTCAGTCTGAGACAAATTCAGAA 927
DB 569 TCGAAGTGGAGCCCGAAAAAAGGCTCAGTCTGAGACCAATTCAGAA 616

RESULT 14
LOCUS      BQ444032      600 bp      mRNA      linear      EST 29-MAY-2002

```

DEFINITION UI-M-EMO-bx1-p-15-0-UI.r1 NIH_BMAP_EMO Mus musculus cDNA clone
IMAGE:5708798 5', mRNA sequence.
ACCESSION B0444032
VERSION B0444032.1 GI:21247144
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 600)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pyx-5.

FEATURES

Source
Location/Qualifiers
1..600
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5708798"
/clone_1lb="NIH_BMAP_EMO"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/note="Organ: brain; Vector: pYX-Asc; Site: 1: EcoR I; Site: 2: Not I. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTGGCTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemlin Chln, Ph.D., program coordinator."

BASE COUNT 145 a 156 c 173 g 125 t 1 others
ORIGIN
Query Match 16.4%; Score 490.2; DB 14; Length 600;
Best Local Similarity 91.4%; Pred. No. 6.2e-58;
Matches 341; Conservative 0; Mismatches 49; Indels 2; Gaps 2;

OY 599 ATTGACACTTGGGATGGGCTCCCTGAGTGGGGGACAGCCCTCGAGACCACTGC 658
DB 251 ATGCGACACTTGTATGGGCTCCCAAGTGGGGACAGCCCTCTGAGACCACTGC 310
OY 659 GGGTCCCGCCATTATGCTGTGTCAGAGGTATTAAGGGGAAAAATATGATGGCCCGG 718
DB 311 GGGTCCCGCCATTATGAGCAGANTCCAGAGGTATTAAGGGGAAAAATGATGGCCCGG 370
OY 719 GCAGACATGTGAGCTGTGGAGTATCTCTTGGCCCTCTGCTGGGGGCTGTGCCCTTT 778
DB 371 GCAGACATGTGAGCTGTGGAGTATCTCTTGGCCCTCTGCTGGGGGCTGTGCCCTTC 430
OY 779 GATGACAGACATCCCGCCAGCTGCTGGAGAAAGGTGAAGGGGCTCTCCACATGCC 838
DB 431 GATGACAGAC-ACCTGGCCAGCTACGTGGAAGGTGAAGGTGCTTCCACATGCC 489
OY 839 CACTTATTCCTCCATATTCAGATTCAGAGCTCTGAGGGAATGATGAGACCCGAA 898
DB 490 CACTTATTCCTCCATATTCAGATTCAGAGCTCTGAGGGAATGATGAGACCCGAG 549
OY 899 AAAAGCTCAGTGTGAGCAATTCAGAAACATCTGTGATACCTAGGCGGGA 950
DB 550 AAAAGCTCAGTGTGAGC-AATTCAGAAACATCTGTGATACCTAGGCGGGA 600

RESULT 15

BF796285 753 bp mRNA linear EST 12-JAN-2001
LOCUS 602258590F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4341838 5',
DEFINITION mRNA sequence.
ACCESSION BF796285
VERSION BF796285.1 GI:12101339
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 753)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Louis Staudt, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:
http://image.llnl.gov
Plate: LLM956 row: c column: 23
High quality sequence stop: 691.

FEATURES

Source
Location/Qualifiers
1..753
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4341838"
/clone_1lb="NIH_MGC_85"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."

BASE COUNT 156 a 212 c 262 g 123 t
ORIGIN
Query Match 16.4%; Score 487.8; DB 12; Length 753;
Best Local Similarity 96.8%; Pred. No. 1.1e-57;
Matches 572; Conservative 0; Mismatches 12; Indels 7; Gaps 7;

```
Db 167 AGCGGGGAAACAGAGGACCAACCGTCCAGAGACCCCGCCGCGGTAGCCAT 226
OY 1003 GCGAGACCTCCATCCAAAGAGAGTGAACCCGACGTCCTAGAGCATGGCATCACT 1062
Db 227 GCGGAGCCTGCTCCATCCAGAGAGAGTGGAGACCCGACGTCCTAGAGCATGGCATCACT 286
OY 1063 GGGCTGCTTCAGGAGACCGCAGAGGCTGCATCCGAGCTGGCAGTGAAGAGAACCA 1122
Db 287 GGGCTGCTTCAGGAGACCGCAGAGGCTGCATCCGAGCTGGCAGTGAAGAGAACCA 346
OY 1123 AGAAAGATGATATATATATCTGCTTTGGATCCGAGAGGCGTATCCACACTGTAGAGA 1182
Db 347 AGAAAGATGATATATATATCTGCTTTGGATCCGAGAGGCGTATCCACACTGTAGAGA 406
OY 1183 CCAGAGACTGCTCCCGGAAATGATGTTGACCCCGGAAAGCGTGGATTCTCCCAT 1242
Db 407 CCAGAGACTGCTCCCGGAAATGATGTTGACCCCGGAAAGCGTGGATTCTCCCAT 466
OY 1243 GCTGAGCCCTGACGAGGAGCGGACCAAGAGAGTCCATGGAAGTCTGAGCATCAC 1302
Db 467 GCTGAGTCTGACGAGGAGCGGAGGAGGAGGAGGAGTCCATGGAAGTCTGAGCATCAC 525
OY 1303 CGATGCGCGGGGCTGCTGCTCCCTGACCCAGACGCGGCTTTGGAGATGCCACAGA 1362
Db 526 CGATGCGCGGGGCTGCTGCTCCCTGACCCAGACGCGGCTTTGGAGATGCCACAGA 585
OY 1363 CAGCCAGAGATCCGTAAGCTCAGT-GGAGCTCCAGGCTGTCTCCAGCCCTTAA 1421
Db 586 CAGCCAGAGATCCGTAAGCTCAGTGGAGACCTCCCGGCTGTCTCTCCAG-CTCTTAA 644
OY 1422 GCAGCCCAAGAGATCCGTAAGCTTCTTTCACCGAGACCGGGGCTGGAGATGAGGCTC 1481
Db 645 GCAG-CCAAGAGATCCGTAAGCTTCTTTCACCGAGACCGGGGCTGGAGATGAGGCTC 702
OY 1482 GAGGGGGGGGCTCCGACTTCCAAAAGCAGAGCTGCTTCGSGGCC 1532
Db 703 GA-GCGGGGGGCTCCCGA-TTCCAAAAGCAGAGCTGCTTCGSGGCC 751
```

Search completed: April 22, 2003, 00:23:38
Job time : 2731.75 secs

GenCore version 5.1.4-P5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 20:17:06 ; Search time 54.0237 Seconds

(Without alignments)
13249.433 Million cell updates/sec

Title: US-10-003-690-3

Perfect score: 2334
Sequence: 1 atgtcgtccggggcgaaga.....ccaacgggacccctctgcc 2334

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	966.6	41.4	2908	US-09-930-181-1	Sequence 1, Appl1
2	951	40.7	3364	US-09-930-181-3	Sequence 3, Appl1
3	206.6	8.9	1647	US-09-101-146-44	Sequence 44, Appl1
4	200.2	8.6	1929	US-09-359-161-4	Sequence 4, Appl1
5	198.6	8.5	1742	US-08-557-006C-38	Sequence 38, Appl1
6	198.6	8.5	2652	US-08-557-006C-39	Sequence 39, Appl1
7	198.6	8.5	2761	US-08-557-006C-24	Sequence 24, Appl1
8	182.6	7.8	1736	US-08-557-006C-37	Sequence 37, Appl1
9	182.6	7.8	1783	US-08-557-006C-36	Sequence 36, Appl1
10	177.8	7.6	2698	US-08-677-298-1	Sequence 1, Appl1
11	175.2	7.5	1747	US-08-557-006C-44	Sequence 44, Appl1
12	145	6.2	213	US-09-930-181-5	Sequence 44, Appl1
13	135.6	5.8	2132	US-09-159-385-3	Sequence 3, Appl1
14	135.6	5.8	2132	US-09-186-277-3	Sequence 3, Appl1
15	130.6	5.6	1428	US-09-159-385-4	Sequence 4, Appl1
16	130.6	5.6	1429	US-09-186-277-4	Sequence 4, Appl1
17	116.6	5.0	1547	US-09-523-849-4	Sequence 4, Appl1
18	116.6	5.0	1544	US-09-734-673-1	Sequence 1, Appl1
19	116.6	5.0	2060	US-09-523-849-1	Sequence 1, Appl1
20	107.2	4.6	1599	US-09-256-465-1	Sequence 1, Appl1
21	107.2	4.6	1599	US-09-167-322-3	Sequence 3, Appl1
22	107	4.6	2610	US-09-212-771-1	Sequence 1, Appl1
23	107	4.6	2610	US-09-091-058-1	Sequence 3, Appl1
24	103	4.4	2549	US-09-467-082-3	Sequence 1, Appl1
25	102.4	4.4	1302	US-08-913-050A-2	Sequence 2, Appl1
26	98.2	4.2	1244	US-08-755-728-1	Sequence 1, Appl1
27	98.2	4.2	1244	US-08-974-655-1	Sequence 1, Appl1

28	98.2	4.2	1244	US-09-283-011-1	Sequence 1, Appl1
29	98	4.2	4162	US-08-459-448A-26	Sequence 26, Appl1
30	98	4.2	4162	US-08-459-595A-26	Sequence 26, Appl1
31	98	4.2	4162	US-08-459-504B-26	Sequence 26, Appl1
32	98	4.2	4162	US-08-459-444-26	Sequence 26, Appl1
33	98	4.2	4162	US-09-547-422-26	Sequence 26, Appl1
34	98	4.2	4165	US-07-951-715A-26	Sequence 26, Appl1
35	97.2	4.2	2104	US-09-313-930-1	Sequence 1, Appl1
36	97	4.2	2061	US-09-800-960-1	Sequence 1, Appl1
37	96.8	4.1	4935	US-08-631-097-3	Sequence 3, Appl1
38	96.8	4.1	5886	US-08-810-712-9	Sequence 9, Appl1
39	96.2	4.1	1282	US-08-878-989-12	Sequence 12, Appl1
40	96.2	4.1	1282	US-09-272-796-12	Sequence 12, Appl1
41	95	4.1	3972	US-09-770-170-1	Sequence 1, Appl1
42	94.2	4.0	1466	US-08-749-902-2	Sequence 2, Appl1
43	93.8	4.0	5228	US-09-428-711A-15	Sequence 15, Appl1
44	93	4.0	1962	US-08-924-183-4	Sequence 4, Appl1
45	93	4.0	1962	US-09-488-364-4	Sequence 4, Appl1

ALIGNMENTS

```

RESULT 1
US-09-930-181-1
; Sequence 1, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 160 101 VI
; CURRENT APPLICATION NUMBER: US/09/930.181
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)..(2112)
US-09-930-181-1

Query Match      41.4%; Score 966.6; DB 4; Length 2908;
Best Local Similarity 70.1%; Pred. No. 3.2e-212;
Matches 1485; Conservative 0; Mismatches 439; Indels 195; Gaps 6;

QY 74 CCCAGCAGCCCAATATGTGGGCCCTATGCGCTGAGAGACGCTGGGCAAGACAGA 133
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 134 CGCAGCAGCGGCACTATGTGGGCCCTACCGCTGAGAGACGCTGGGCAAGGCGACA 193

QY 134 CAGGCGGTAAATCGGGGTCACATGACATCAGCGGTGAGAGAGTGGCCATCAAGATCG 193
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 194 CAGGCTGGTGAACCTGGGGGTCTACCTGCTGCAGAGAGTGGCCATCAAGATCG 253

QY 194 TGAACCGGAGAACCTGTGCGAGTCGTGATGAAGTGGAGCGGAGATCGCATCC 253
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 254 TCAACCGTGAAGAACCTGACGAGTCGTGATGAAGTGGAGCGGAGATCGCATCC 313

QY 254 TGAAGCTCATGAGAACCCACCATGTCTCAAGCTCCAGAGCTTACGAGAAAGAAAT 313
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 314 TGAAGCTCATGAGAACCCACCATGTCTCAAGCTCCAGAGCTTACGAGAAAGAAAT 373

QY 314 ATTGTACCTGGTTCGTGAGACGCTCGGGGGGTGAGCTATTCAGTACCGGTAAAGA 373
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 374 ATTGTACCTGGTTCGTGAGACGCTCGGGGGGTGAGCTATTCAGTACCGGTGAGAA 433

QY 374 AGGGAAGACTGAGCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 433
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 434 AGGGAAGACTGAGCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 493

QY 434 ACTTCTGCACAGCTACTCTCATCTGCGACAGAGACCTTAAAGCCGAGAACTGCTTGG 493

```

Db	494	ACTTCTGCCACACCACTCCATATGCGCACAGAGGATCTGAACACTGAACACTCCATGCTCTG	553
OY	494	ATGAGAAAAACAATCCGATTCGACACTTCGGCATATGGCGTCCCTCGAGCGGGGAC	553
Db	554	ACGAGAAACAACATCCGATCGACACTTGGCATATGGCGTCCCTCGAGCGGGGAC	613
OY	554	GCCTCTGAGACACGACTGCGGGTCCCCCATATATGCTGTCCAGAGGTATTAAGGGG	613
Db	614	GCCTGTGGAGACAGCTGTGGGTCCCCCACTACGCTTGCCCCGAGGTACCGGGGG	673
OY	614	AAAAATATATGCCCCCGGGGACACATGTGGAGCTGTGGAGTACTCTTCCCTCC	673
Db	674	AGAGATATGACGGCCGGAAAGCGGACGTGTGGACTGTGGGCTCATCTGTTCGCTTC	733
OY	674	TCGTGGGGGCTTGCCCTTGTATACACAACAATCCGCGACGTGCTGGAGAGTGAAC	733
Db	734	TGCTGGGGGCTTGCCCTTGTAGCATATGACACTTCCGACACTGCTGGAGAGGTGAAC	793
OY	734	GGGGGCTTTCACATGCCCCCATTCTCTCAGATTGCGCAGAGCTCCTGAGGGGA	793
Db	794	GGGGGCTTTCACATGCCCCCATTCTCTCAGATTGCGCAGAGCTCCTGAGGGGA	853
OY	794	TGATCAAGTGGAGCCCGAAAAAGGTCAGTCTGGAGCAATTCAGAAACATCTCTGT	853
Db	854	TGATCAAGTGGAGCCCGAACCGCCGCTCACGCTAGACCAATTCAGAAACATATGTT	913
OY	854	ACCTAGCGGGGAACACAGACCAACCCGTGCTCGAGGCCAGCCCTGAGCCGGTAG	913
Db	914	ATATAGGGGGCAAGATAGACCCGGAAC-----AGAGACCCCATTCCTCGAAGTGC	967
OY	914	CCATGGGGAGCCCTGCCATCCACGGAAGCTGGAGCCCGACGTCTTAAGAGCATGGCAT	973
Db	968	AGATCCGCTCGTCCGACGCTGGAGACATCGACCCCGACGTGCTGGACACATGCACT	1023
OY	974	CACCTGGGCTCTTCAGAGGACCGGAGAGGCTGCATCGGAGCTGCGCATGAGGAGAGA	1033
Db	1028	CATGGGCTCTTCGAGAGCCGACAAAGCTGTGACAGACTGCTGTCCAGAGAGAGA	1083
OY	1034	ACCAAGAAAAGATGATATATATATCTGCTTTGGATCGGAGAGCGGTATCCACGCTGTG	1093
Db	1088	ACCAGAGAGAATGATTTACTTCTCTCTCGGACCGGAAAAAGTATACCCGACCAAG	1143
OY	1094	AGGACACGAGACCTGCGCTCCCGGAAATGATGTGACCCCCCGGAAAGGTGGAGTTTC	1153
Db	1148	AGGATATAGGACCTGCCCCCCCCGGAACGATAGACCTCCCGGAAAGGTGGAGTTTC	1203
OY	1154	CCATGCTGAGCCGCTCACGGGAGCGGACGACGACGAGCGGAAGTCCATGAACTCGAGCA	1213
Db	1208	CGATGCTGAACCGGACGCGCAAGGGGGGCGCGAGAACGCAATGCATGAGGTGCTCAGG	1263
OY	1214	TCACCCATCCGGGGGTGTGTGCTCCCTGTACCCACCCGACGGGCTTGGAGATGGCCC	1273
Db	1268	TGAC-----GGACGGGGGTTCCTCCCGGTGCTGGCGCGGGCCATTGAGATGGCCC	1313
OY	1274	AGCACACGAGAGATCCGATGCTGATGAGGCTCCACGCGGTCTGCTCCACGCTC	1333
Db	1319	AGCAGGCGCAGAGTCTCGGTCCATCAGCGGTGCTCTCAGGCTTTCACACAGCCAC	1373
OY	1334	TAAAGCACCCCAAGAGCTCGGGTCTTTTCCTTTACCGGAGCGGGGCTGGAAGTAGG	1393
Db	1379	TCAGAGGCCCCG-----	1393
OY	1394	CTCGAGGGGGGCTCCCGCACTTCCAAAAGCAGAGCGTGCCTTCTCGGGGCCCAAGG	1453
Db	1392	-----	1393
OY	1454	GTGGGGGCGCCGGGAGACAGCCCGCGCCCACTGCGCGCTCCACACCCCTGCGGGCC	1513
Db	1392	-----GATACCCCTACCCCTTCACCAAGGGGAGTGTCCCTCCCC-----	1433
OY	1514	CCCCAGGCTCCCGCGCTCTCTGTGGGGGAAACCCCTTGCACCTGGGCTGTGACACAGCCCC	1573

Db: 1432 -----ACCCCAAGGGGAGAACCTGTCCACAGGCCAA 1462

QY 1574 GGGCCAGTCCCAACCGGAGCCCCGGGAGACAACACACCACCCCGGCGGGGTGGCTCG 1633

Db 1463 AGGAGAGCCCGGTGGACGGCCCAACCCCAAGCCCGCTGACACCCC-----AGCGTCG 1516

QY 1634 GGGGAGACCCGCTGGAGGAGTGTCTCAACTGCATCCGACACAGCTTCTGGGGTCCCTC 1693

Db 1517 GAGGGTGGCCCTGGAGGGCGCGGCTCAACTGCATCAAGAACAGCTTCTTGGGCTCACCCC 1576

QY 1694 GCTTTCACGGCGGAGATGCAATGCACTCCCTACCGCTGAGAGATGTCCAGCTTGAACGCGAG 1753

Db 1577 GCTTTCACGGCGGAGAACTGCAAGTGTCCGACGCGCGAGAGATGTCCAACTGACACAG 1636

QY 1754 ACTCTCCCCCGAGCTGGGCAAAACGCTCTCGTGTTCGGGAACCTTCATCTCTTGGACAAAG 1813

Db 1637 ACTGTGCCCCAGAGCTGGGCAAAAGAACTCTGTTGGGAACCTTCATCAGCTTGGAGAGG 1696

QY 1814 AAGAACAAATATTTCCGTGTCTAAAGAGAACAAACCTGTCACAGATGATCAAGAGACATCG 1873

Db 1697 AAGAGACAGATCTTCTGTGTTCATCAAAAGACAAACCTCTGAGCTCTCATCAAGGCTGACATCG 1756

QY 1874 TCCATGCTTTCTGTGATCCCAAGCCTGATGACAGTGTGTCTGTCACAGACAGCTTCA 1933

Db 1757 TCGACGCTTCTGTGTGATTTCCAGTCTCAGCCACAGCGTCAATCTCCCAAGAGCTTCC 1816

QY 1934 GGGCGGAGTACAAAGGCCAGTGGCGGCCCTCGTCTTCCAAAGCCCGCTCCGCTTCAGG 1993

Db 1817 GGGCGGAGTACAAAGGCCAGCGGGGGGCGCACGCCGTGTTCAGAAAGCCGCTCAAGTTCAGG 1876

QY 1994 TGGACATCAGCTCTCTGAGAGGTCCAGAAGCCCTCCCGGAGCGGAGCGGAGGCTG 2053

Db 1877 TTGATATCTCACTACACGGAGGT-----GGGGAGGCGCAAGAGAGA 1918

QY 2054 GTGGCATCTACTCCGTCACCTTCATCTCATCTCGGGTCCAGACCGCTGCTCAAGCGAG 2113

Db 1919 ACGGCATCTACTCCGTCACCTTCATCTCATCTCGGGTCCAGACCGCTGCTCAAGCGAG 1978

QY 2114 TGTGTGAGACCATCCAGGCGACAGCTCTGAGCACTATGACACCAACCCCTCGCTGAGGCGCC 2173

Db 1979 TGTGTGAGACCATCCAGGCGCGCTGAGCACTATGACACCAACCCCTCGCTGAGGCGCCAGCT 2038

QY 2174 TGGCAGAGAGAGAAGACGG 2192

Db 2039 TGTCAAGACACCTACTG 2057

RESULT 2

US-09-930-181-3

; Sequence 3, Application US/09930181

; Patent No. 6455292

; GENERAL INFORMATION:

; APPLICANT: Origene Technologies

; TITLE OF INVENTION: Full-length Serine Protein Kinase In Brain and Pancreas

; FILE REFERENCE: 160 101 VI

; CURRENT APPLICATION NUMBER: US/09/930,181

; CURRENT FILING DATE: 2001-08-16

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 3

; LENGTH: 3364

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (482)..(2239)

US-09-930-181-3

Query Match 40.7%; Score 951; DB 4; Length 3364;

Best Local Similarity 69.9%; Pred. No. 1,2e-208;

Matches 1484; Conservative 0; Mismatches 440; Indels 199; Gaps

74 CCCAGCAGCCCAATATATGTGGGCCCTATCGCTGAGAGAAAGACGCTGGGCAAGAGCAGCA 133

Db 257 CGGAGCGAGCGAGTATGTTGGCCCTACCGCGTGAGAAAGCGTGGGCAAGGGGCA 316
 QY 134 CAGGCGTGGTTAACTGGGGTCCATGTCATCAGGGGTGAGAAAGTCCGATCAAGATG 193
 Db 317 CAGGTCTGGAGAGCTGGGGTTCACCTCCGACCTGCGAGAAAGTGGCCATCAAGATG 376
 QY 194 TGAACGGGGAAGAGTCTGGAGTGGTCTGATGAAAGTGGAGCGGAGATGCCATCC 253
 Db 377 TCAACCGTGAAGAGCTCAGGAGTGGGTGATGAAAGTGGAGCGGAGATGCCATCC 436
 QY 254 TGAAGCTCATGAAACCCCATATGCTTCAAGCTCCAGAGCTCCAGAGAAACAAGAAAT 313
 Db 437 TGAAGCTCATGAAACCCCATATGCTTCAAGCTCCAGAGCTCCAGAGAAACAAGAAAT 496
 QY 314 ATTT-----GTACCTGGTTCTGGAGACGTCGCGGGGGTGAAGCTATTCGACTACCTGTA 369
 Db 497 ATTTGATGATACCTGGTCTGGAGACGTCGCGGGGGTGAAGCTATTCGACTACCTGTA 556
 QY 370 AAGAAGGGGAGACTGACGCCCAAGAGGCGCGAAAGTTCTTCGCGAGATTTGCTGCG 429
 Db 557 AAGAAGGGGAGGCTGAGCGCTTAAGAGGCTGGAAGTTCTTCGCGAGATTCATCTGCG 616
 QY 430 CTGGACTTCCGACAGCTACTCATGTCGACAGAGACCTTAAGCCCGAGAACTGCTT 489
 Db 617 CTGGACTTCCGACAGCTACTCATGTCGACAGAGACCTTAAGCCCGAGAACTGCTT 676
 QY 490 TTGGTGAAGAAAAACAATCCGATTCGAGACTTCGCGATGGCTCCCTGCAAGTGGG 549
 Db 677 CTGGAGAGAAAGAAACAATCCGATTCGAGACTTCGCGATGGCTCCCTGCAAGTGGG 736
 QY 550 GACAGCTCTGAGACCAAGCTGGGGTCCCGCATTAATGCTGTCCAGAGGTATTAAG 609
 Db 737 GACAGCTCTGAGACCAAGCTGGGGTCCCGCATTAATGCTGTCCAGAGGTATTAAG 796
 QY 610 GGGGAAAAATATGATGGCGCGCGGAGACAAATGAGAGCTGAGTATCTCTGGCG 669
 Db 797 GGGGAAAAATATGATGGCGCGCGGAGACAAATGAGAGCTGAGTATCTCTGGCG 856
 QY 670 CTGCTCTGGGGGCTGCTGCTTGTATGACACAACCTCCGACAGCTGCTGAGAGATG 729
 Db 857 TTGCTGTGGGGGCTGCTGCTTGTATGACACAACCTCCGACAGCTGCTGAGAGATG 916
 QY 730 AAACGGGCGCTTCCACAGCCCGCACTTATCTCTCAGATGCGCAGAGCTCTGAG 789
 Db 917 AAAGGGGCGGCTTCCACAGCCCGCACTTATCTCTCAGATGCGCAGAGCTCTGAG 976
 QY 790 GGAATGATCGAAGTGAAGCCCGGAAAAAGGCTCATGAGCAATTCAGAAACATCT 849
 Db 977 GGAATGATCGAAGTGAAGCCCGGAAAAAGGCTCATGAGCAATTCAGAAACATCT 1036
 QY 850 TGTACTAGCGGGGAAACACAGCCAGACCCGCTGAGAGCCAGCCCTGGCGCG 909
 Db 1037 TGTACTAGCGGGGAAACACAGCCAGACCCGCTGAGAGCCAGCCCTGGCGCG 1090
 QY 910 GTAGCCATGCGGAGCTGCTCCATCAAGCGAGAGCTGGAACCCGAGCTCTGAGAGCATG 969
 Db 1091 GTAGCCATGCGGAGCTGCTCCATCAAGCGAGAGCTGGAACCCGAGCTCTGAGAGCATG 1150
 QY 970 GCATCATGCGGCTCTTCAGGAGCGCGAGAGGCTGATCCGAGCTCGGAGTGAAG 1029
 Db 1151 GCATCATGCGGCTCTTCAGGAGCGCGAGAGGCTGATCCGAGCTCGGAGTGAAG 1210
 QY 1030 GAGAACCAAGAAAGATGATATTTATCTGTTTGAATCGAAGAGAGGCTATCCAC 1089
 Db 1211 GAGAACCAAGAGATGATATTTATCTGTTTGAATCGAAGAGAGGCTATCCAC 1270
 QY 1090 TGTAGGACCAAGAGCTGCTCCCGAATGATGTTGAACCCCGGAGAGGCTGAT 1149
 Db 1271 TGTAGGACCAAGAGCTGCTCCCGAATGATGTTGAACCCCGGAGAGGCTGAT 1330
 QY 1150 TCTCCCATGCTGAGACCGCTCAGGAGGAGGCGGAGCGGAGAACTCCATGAAAGCTCTG 1209

Db 1331 TCCCCCATGCTGAACCGGACGCGAGCGGCGCCAGACGCAAAATTCATGAGGTCTC 1390
 QY 1210 AGCATCAACCATGCGGGGGTGTGTGCTCCCTGTACCAACCGAGGCGCTTGGAGATG 1269
 Db 1391 AGGTTGAC-----GAGCGGGGGCTCCCGGGTGGCTGGCGGGCCATTGATG 1441
 QY 1270 GCCCAGCAACCAAGAGATCCGTTAGCTGAGTGAAGCTTCAGAGGTTCTGCTCCAGC 1329
 Db 1442 GCCCAGCAACCAAGAGATCCGTTAGCTGAGTGAAGCTTCAGAGGTTCTGCTCCAGC 1501
 QY 1330 CCTCTAAGCAAGCCCAAGAGATCCGTTCTTTCTTTTACCGGAGCCGGGGCTGAGAT 1389
 Db 1502 CCACTCAGCAGCC----- 1514
 QY 1390 GAGGCTGAGAGCGGGGCTCCCGACTTCAAAAGCAGAGCTGCTTTCGGGGCCCC 1449
 Db 1515 ----- 1514
 QY 1450 AGGGGTGGGGCGCGGGGAGACGCCCGCCCGCACTGCCCCCTCCACACCCCTGCC 1509
 Db 1515 -----CCGGGTGACCCCTCACCCCTCACAAAGGGGAGTCCCTCCCC----- 1558
 QY 1510 GGGCCCCCAGGCTCCCGGCTCCTGCGGGGAGCCCGCTTGCACTGGCTCTGACAG 1569
 Db 1559 -----ACCCCAAGGGAGACACTGTCCACAG 1585
 QY 1570 CCGCGGCGACGTCACCGGAGCCCGGGAGACACCAACCCCGCCGCGCGGCGGTGAC 1629
 Db 1586 CCAAGGAGAGAGCCGGGTGGAGGCCCAAGCCCAAGGCCCGGCTGCAAGCCCC-----AGC 1639
 QY 1630 GTGGGGGAGCGCGCTGAGAGATGCTCTCAACTCATCCGCAACAGTCTCTGGGCTCC 1689
 Db 1640 GTGGGGGAGCGCGCTGAGAGATGCTCTCAACTCATCCGCAACAGTCTCTGGGCTCC 1699
 QY 1690 CTTGCTTCAACCGGCAAGATGAGTCCCTACCGCTGAGAGATGCTCCAGTTCAGC 1749
 Db 1700 CCGCGCTTCAACCGGCAAGATGAGTCCCTACCGCTGAGAGATGCTCCAGTTCAGC 1759
 QY 1750 CCAAGTCTCCCGCGGAGCTGGCAAAAGCTCTGTTGCGGAACTTCATCTCTTGGAC 1809
 Db 1760 CCAAGTCTCCCGCGGAGCTGGCAAAAGCTCTGTTGCGGAACTTCATCTCTTGGAC 1819
 QY 1810 AAAGAAACAATATTTCTGCTTAAGAGCAAACTCTGAGCAGCATCAAGCAGAC 1869
 Db 1820 AAAGAAACAATATTTCTGCTTAAGAGCAAACTCTGAGCAGCATCAAGCAGAC 1879
 QY 1870 ATGCTCATGCTTCTGTGATCCCAAGCTGAGTCAAGTGTCTGACAGACAGC 1929
 Db 1880 ATGCTCATGCTTCTGTGATCCCAAGCTGAGTCAAGTGTCTGACAGACAGC 1939
 QY 1930 TTCAGGCGCGAGTACAAGGCGAGTGGCGCCCTCTCTCTTCAAAAGCCCTCGCTTC 1989
 Db 1940 TTCAGGCGCGAGTACAAGGCGAGTGGCGCCCTCTCTCTTCAAAAGCCCTCGCTTC 1999
 QY 1990 CAGGTGACATGACTCTCTGAGGGTTCAGAGCCCTCCCGCAGAGGAGCGGACGGA 2049
 Db 2000 CAGGTGACATGACTCTCTGAGGGTTCAGAGCCCTCCCGCAGAGGAGCGGACGGA 2041
 QY 2050 GGTGGTGGCATCTACTCGTACCTTCACTCATCTGCGGGCCGAGCGTGGTTCAAG 2109
 Db 2042 GAGAACCGCATCTACTCGTACCTTCACTCATCTGCGGGCCGAGCGTGGTTCAAG 2101
 QY 2110 CGAGTGGTGGAGACCATCAGGACAGTCTCTGAGACATGACAGCCCTCGCTGAG 2169
 Db 2102 AGGGTGGTGGAGACCATCAGGACAGTCTCTGAGACATGACAGCCCTCGCTGAG 2161
 QY 2170 GCGCTGACAGAGAGAGAGCG 2192
 Db 2162 CACTTGTACAGACCACTAAGT 2184

RESULT 3
 US-09-101-146-44

QY 319 TACCTGTTCTGAGACGCTCTCGGGGGTACCTATTGACTACTCTGTAAGAAGGG 378
 DB 309 TATGTTGTGATGGAGTATGGAATCTGGGAGATTATGATTTGATTTGATGAGAGGGC 368
 QY 379 AGACTGACGCCCAAGAGGCGGAGAGCTCTTCCGCGGATTTGCTGCGCTGAGCTTC 438
 DB 369 AGATTGACGAGATGAGAGCTCTTACTTTTTCAGCAGATATTTCTGCTGAGATAC 428
 QY 439 TGGCAGAGCTACTCTCATCTGCGACAGAGACTAAAGCCGGAACCTCTTTTGATGAG 498
 DB 429 TGGCATAGAAACATGATGTTTATAGAGACTTAAGCTGAAGCTCTCTCTGAGCTCC 488
 QY 499 AAAAAACATCTCCGATTCGAGACTTGGGCTGCGCTCTGAGAGTGGGGACAGCTTC 558
 DB 489 AATGAGATGATGAGATGAGATTTGTTGATGACATATCATGCGGATGCTATTTT 548
 QY 559 CTGAGAGACGCTGCGGGGCTCCCATTTATGCTGCTGAGAGGATTAAGGGGAAAAA 618
 DB 549 CTGAGAGACGCTGCGGGGCTCCCATTTATGCTGCTGAGAGGATTAAGGGGAAAAA 608
 QY 619 TATGATGCGCGCGGAGACATGTGAGCTGTGAGATCATCTTCTGCGCTGCTGCTG 678
 DB 609 TATGCTGCGCGCTGAGATGATGATGAGCTGTGATGATTTATCTTATGCTCTCTGCT 668
 QY 679 GGGGCTCTGCGCTTGTGATGAGACACCTCCGCACTGCTGAGAGAGTGAAGGGGC 738
 DB 669 GGCACCTCTGCGCTTGTGATGAGACACCTCCGCACTGCTTGTGAGAGAGTGAAG 728
 QY 739 GTCCTGACATGCGCGGCTCTCATCTCTGACATGCTGAGAGCTCTGAGGGGATGATC 798
 DB 729 ATATATATCTGCGCGGCTCTCATCTCTGACATGCTGAGAGGATGATGATGATGATC 788
 QY 799 GAAGTGAAGCCGCAAAAGAGCTGCTGAGAGCAATTCAGAAACATCTGTGATCCTA 858
 DB 789 ATATGATGACCAATGAGAGCAATGATCTGATGATGCTGATGCTGATGCTGATGCTG 848
 QY 859 G 859
 DB 849 G 849

RESULT 5
 US-08-557-006C-38
 ; Sequence 38, Application US/08557006C
 ; Patent No. 6258547
 ; GENERAL INFORMATION:
 ; APPLICANT: Berl, Rajindar K.
 ; APPLICANT: Carling, David
 ; APPLICANT: Forster, Robert A.
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
 ; FILE REFERENCE: NGAP/PM37588/UST
 ; CURRENT APPLICATION NUMBER: US/08/557,006C
 ; CURRENT FILING DATE: 1996-03-06
 ; PRIOR APPLICATION NUMBER: PCT/GB94/01093
 ; PRIOR FILING DATE: 1994-05-20
 ; PRIOR APPLICATION NUMBER: GB 9310489.1
 ; PRIOR FILING DATE: 1993-05-21
 ; PRIOR APPLICATION NUMBER: GB 9318010.7
 ; PRIOR FILING DATE: 1993-08-31
 ; NUMBER OF SEQ ID NOS: 44
 ; SOFTWARE: Patentln Ver. 2.1
 ; SEQ ID NO 38
 ; LENGTH: 1742
 ; TYPE: DNA
 ; ORGANISM: Human AMP protein kinase
 ; US-08-557-006C-38

Query Match 8.5%; Score 198.6; DB 4; Length 1742;
 Best Local Similarity 54.6%; Pred. No. 1.7e-36;
 Matches 418; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

QY 92 TGGGCCCCATATGCGCTGAGAGAGAGCTGGGCAAGAGACAGAGGCTGTGTTAACTCG 151

DB 38 TCGGACACTACGCTGCTGGGGACACCTGGGCTGCGGACCTTGGCAAGTGAAGATTG 97
 QY 152 GGGTCCACTGATCATTCACGGGTGAGAGTCCGACATCAAGATGTAACCGGGAGAGCTGT 211
 DB 98 GAGAACATCAATTTGACAGGGCATTAAGTGGCATGTTAATTTAATGACAGAAAGATTG 157
 QY 212 CGGAGT---CGGTCTGATGATGAGTGAAGGGGAGATGCGCATCTGAGACTCATGAA 268
 DB 158 GCAGTTAGATGTTGTTGGAATAAATAAGGAAATTCATAAACTTCTTCTGCTC 217
 QY 269 ACCGACATGCTCCAGCTGACAGAGCTGATGAGAGAACAAATTTGATACCTGCTTC 328
 DB 218 ATCTCATATATATCAAACTTACCAAGTATGATGATGATGATGATGATGATGATGAT 277
 QY 329 TGGAGACGCTCTGCGGGGCTGAGCTATTCGATCTGATGATGATGATGATGATGATGAT 388
 DB 278 TGGATATATGCTGCGGAGTGAATTTGATGATGATGATGATGATGATGATGATGATGAT 337
 QY 389 CCAAGAGGCGGCAAACTTCTTCCGCAAGATTTGCTGCGCTGAGCTTCTGCAAGCT 448
 DB 338 AGGTGAGAGCTGCGCGGCTCTTCCAGAGATTTCTGCTGCGGCTGCTGCTGCTGCTG 397
 QY 449 ACTCATCTGCGACAGAGCTGATGAGAGGCGGAGAGAGCTGCTTGGATGAGAAAAACA 508
 DB 398 ACATGCTGCTGCGAGAGGCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 457
 QY 509 TCCGATATGAGAGCTGCGGATGAGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 568
 DB 458 CTAAAGATGAGCTGAGCTGAGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 517
 QY 569 GCTGCGGGTCCCCATTTATGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 628
 DB 518 GCTGCTGATGAGCAAAATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 577
 QY 629 GCGGAGAGAGATGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 688
 DB 578 CTGAGGTTGATGATGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 637
 QY 689 CTTTGTATGAG 748
 DB 638 CTTTGTATGAG 697
 QY 749 TGGCCCACTTCAATCTTCCAGATTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 808
 DB 698 TCCCGAGATGATGAG 757
 QY 809 CCGAATAAAGAGCTGAG 853
 DB 758 CTTTGAAGGAG 802

RESULT 6
 US-08-557-006C-39
 ; Sequence 39, Application US/08557006C
 ; Patent No. 6258547
 ; GENERAL INFORMATION:
 ; APPLICANT: Berl, Rajindar K.
 ; APPLICANT: Carling, David
 ; APPLICANT: Forster, Robert A.
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
 ; FILE REFERENCE: NGAP/PM37588/UST
 ; CURRENT APPLICATION NUMBER: US/08/557,006C
 ; CURRENT FILING DATE: 1996-03-06
 ; PRIOR APPLICATION NUMBER: PCT/GB94/01093
 ; PRIOR FILING DATE: 1994-05-20
 ; PRIOR APPLICATION NUMBER: GB 9310489.1
 ; PRIOR FILING DATE: 1993-05-21
 ; PRIOR APPLICATION NUMBER: GB 9318010.7
 ; PRIOR FILING DATE: 1993-08-31
 ; NUMBER OF SEQ ID NOS: 44
 ; SOFTWARE: Patentln Ver. 2.1
 ; SEQ ID NO 39

```

; LENGTH: 2652
; TYPE: DNA
; ORGANISM: Rat
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)-(1747)
; OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -
; OTHER INFORMATION: Fragment begins at nucleotide 24 and ends with
; OTHER INFORMATION: nucleotide 1765
; US-08-557-006C-39

```

```

Query Match      8.5%; Score 198.6; DB 4; Length 2652;
Best Local Similarity 54.6%; Pred. No. 1.9e-36;
Matches 418; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

```

```

QY 92 TGGGCCCCATCGGCTGAGAAAGAGAGCTGGGCAAGACAGACAGGCGCTGTTAACTCG 151
DB 45 TCGGACACTACGCTGGGGGACACCCCTGGCGCGGACCTCGGCAAGAGGATTTG 104
QY 152 GGGTCCACATGACAGCGGCTCAGAGGTCCCATCAAGATCGTGAACCGGAGAGACTGT 211
DB 105 GAGAACATCAATGACAGCGCATTAAGTGGCAGTTAAGTCTTAATAGACAGAGATTC 164
QY 212 CGGAGT---CGGTGCTGATGAAGTGGAGCGGAGATCGCATCTCTGAAGCTCATCGAAC 268
DB 165 GCAGTTAGATGTTGTTGGAATAAATTAACGAGAAATTCCTTAACCTCTTCGTC 224
QY 269 ACCACATGCTCTCAAGCTCCAGAGCTCTACAGAACAGAAATATTGTACTGTTTC 328
DB 225 ATCTCATATTATCAAACTCTACCAAGTGTACAGACTCCACAGACTTTTATATGTTAA 284
QY 329 TGGAGCAGCTCTCGGGGGGTGAGACTTTCGACTACCTGTGTAAAGAGGAGAGCTGACGC 388
DB 285 TGGAAATATGTTGTGGAGTGAATGTTGCTACTCATCTGTAACACGAGGGGTTGAAG 344
QY 389 CCAAGGAGCCCGAAAGTCTCCGCAAGATTGTGTGCGCTGGAGCTTGCCACAGCT 448
DB 345 AGGTGGAAGCTCCCGGCTCTTCCAGCAGATTGTGTGCGCTGGAGCTTGCCACAGCT 404
QY 449 ACTCATCTGCGACAGAGACTTAAAGCCGAGAACTGCTTTGGATGAGAAAAACA 508
DB 405 ACATGTTGTCACAGAGGACTGAGAGCAAGACGTGTTCTGAGCCGACAGATGATG 464
QY 509 TCCGCATGAGACTTGGCATGGCTCCCTGCAAGTGGGGGACAGCTCTCGAGACCA 568
DB 465 CTAAAGATAGCTGACTCGGACTCTCTAATATGATGTAGATGTGAATTTCTACGAACTA 524
QY 569 GCTGCGGGTCCCCCATTAATGCGTGCAGAGGTGAATTAAGGGGAAAAAATATGATGACC 628
DB 525 GCTGTGAGATGCCAAATTAATGACAGACCGGAGGTCAATCTCAGAAAGGCTGTATGCGGGTGC 584
QY 629 GCCGGCAGACATGTGAGCTGTGAGTCAATCTTCCGCTGCTGCGGGGCTGTGC 688
DB 585 CTGAGGTTGATATCTGAGCTGTGAGTGTATCTGTATGCGCTTCTGTGTGACACCTTCC 644
QY 689 CCTTTATGACGACAACTCCGCAAGCTGTGAGAAAGTGAAGGGGCGCTTTCCACA 748
DB 645 CGTTACACGATGAGCAAGTCTCTACGCTCTTAAAGATCCGAGGGGTGTGTTCTACA 704
QY 749 TGCCTCCACTTCAATCTCCAGATTGCCAGAGCCCTCGAGGGGAATATGCAAGTGGAGC 808
DB 705 TCCCGGAGTATCTCAACCGTTCTATTGCCACTGTGCTGATGACATGTCGAGGTGAC 764
QY 809 CCGAAAAAAGGCTCAGTCTGAGCAAAATTCAGAAACATCTCTGTT 853
DB 765 CCTTGAAGGAGCAACATCAAGACATAGAGAGCATGATGTT 809

```

```

RESULT 7
US-08-557-006C-24
; Sequence 24, Application us/08557006C
; Patent No. 6258547
; GENERAL INFORMATION:

```

```

; APPLICANT: Berl, Rajindar K.
; APPLICANT: Carling, David
; APPLICANT: Forster, Robert A.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
; FILE REFERENCE: NCAP/PM137588/US*
; CURRENT APPLICATION NUMBER: US/08/557, 006C
; CURRENT FILING DATE: 1996-03-06
; PRIOR APPLICATION NUMBER: PCT/GB94/01093
; PRIOR FILING DATE: 1994-05-20
; PRIOR APPLICATION NUMBER: GB 9310489.1
; PRIOR FILING DATE: 1993-05-21
; PRIOR APPLICATION NUMBER: GB 9318010.7
; PRIOR FILING DATE: 1993-08-31
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 24
; LENGTH: 2761
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: cDNA of rat
; OTHER INFORMATION: Liver AMP protein kinase
; US-08-557-006C-24

```

```

Query Match      8.5%; Score 198.6; DB 4; Length 2761;
Best Local Similarity 54.6%; Pred. No. 1.9e-36;
Matches 418; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

```

```

QY 92 TGGGCCCCATCGGCTGAGAAAGAGAGCTGGGCAAGACAGACAGGCGCTGTTAACTCG 151
DB 61 TCGGACACTACGCTGGGGGACACCCCTGGCGCGGACCTCGGCAAGTGAAGTTG 120
QY 152 GGGTCCACATGACAGCGGCTCAGAGGTCCCATCAAGATCGTGAACCGGAGAGCTGT 211
DB 121 GAGAACATCAATGACAGGCAATTAAGTGGCAGTTAAGTCTTAATAGACAGAGATTC 180
QY 212 CGGAGT---CGGTGCTGATGAAGTGGAGCGGAGATCGCATCTCTGAAGCTCATCGAAC 268
DB 181 GCAGTTAGATGTTGTTGGAATAAATTAACGAGAAATTCCTTAACCTTCTTCGTC 240
QY 269 ACCACATGCTCTCAAGCTCCAGAGCTCTACAGAGCTCTACAGAACAGAAATATTGTACTGTTTC 328
DB 241 ATCTCATATTATCAAACTCTACCAAGTGAATGACACTCCACAGACTTTTATGTTAA 300
QY 329 TGGAGCAGCTCTCGGGGGGTGAGCTATTGACTACTGTTAAAGAAAGGAGAGCTGACGC 388
DB 301 TGGAAATATGTTGTGAGGAGTAATGTTGTGACTACATCTTAACAGAGGAGGTTGAG 360
QY 389 CCAAGGAGCCGAAAGTCTTCCGCCAGATGTGTCTGCGCTGAGACTTGTGCCACAGCT 448
DB 361 AGGTGGAAGCTCCCGGCTCTTCCAGCAGATGTTGTCTGCGCTGAGACTGTCCACAGGC 420
QY 449 ACTCATCTGCGACAGAGACTTAAAGCCGAGAACTGCTTTTGGATGAGAAAAACA 508
DB 421 ACATGTTGTCACAGAGGACTGAAGCCAGAAACGTTGTCGAGAGCCAGATGATG 480
QY 509 TCCGATTCAGACTTCGCGATGCGCTCCCTGCAAGGTGAGGGGAGAGAGCTCTGAGACCA 568
DB 481 CTAAAGTACTGACTCTCGGACTCTCTAATATGATGATGATGATGATGATGATGATGATG 540
QY 569 GCTGCGGGTCCCCCATTAATGCGTGTCCAGAGGTGATTAAGGGGAAAAATATGATGAGCC 628
DB 541 GCTGTGAGATGCCAAATTAATGAGCAGCCGAGGTTCAATCTCAGGAAGGCTGTATGCGGGTGC 600
QY 629 GCCGGCAGACATGAGAGCTGTGAGTCAATCTCTTCCGCGCTGCTGCTGAGGGGCTGTGC 688
DB 601 CTGAGGTTATATCTGAGACTGTGAGTGTATCTGTATGCTCTTCTGTCGACACCTTCC 660
QY 689 CCTTTGATGACGACAACTCCCGCAGCTGCTGAGAGAGTGAAGAGGGGCGCTTCCACA 748
DB 661 CGTTGACGATGAGCAGACGTGCTACGCTCTTTAAAGAAATCCGAGGGGAGTGTGTTCTACA 720
QY 749 TCCCCCATTTCAATCTCCAGATTGGCAGAGCCTCTGAGGGGAATGATGAAGTGGAGC 808

```

Db 721 TCCGGAGATCTACACGCTCTATGCGCCTCTGCTATGACATCTGCGAGGTGACC 780
 Oy 809 CCGAAGAAAGGCTCAGCTTGAGCAAAATTCAGAAACATCCTTGGT 853
 Db 781 CCTGAAGCGAGCAACTATCAAGACATACGAGACATGATGATG 825

RESULT 8

US-08-557-006C-37
 ; Sequence 37, Application US/08557006C
 ; Patent No. 6258547
 ; GENERAL INFORMATION:
 ; APPLICANT: Berl, Rajindar K.
 ; APPLICANT: Carling, David
 ; APPLICANT: Forster, Robert A.
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
 ; FILE REFERENCE: NGAP/PH37588/OST
 ; CURRENT APPLICATION NUMBER: US/08/557,006C
 ; CURRENT FILING DATE: 1996-03-06
 ; PRIOR APPLICATION NUMBER: PCT/GB94/01093
 ; PRIOR FILING DATE: 1994-05-20
 ; PRIOR APPLICATION NUMBER: GB 9310489.1
 ; PRIOR FILING DATE: 1993-05-21
 ; PRIOR APPLICATION NUMBER: GB 9318010.7
 ; PRIOR FILING DATE: 1993-08-31
 ; NUMBER OF SEQ ID NOS: 44
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 37
 ; LENGTH: 1736
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Rat liver AMP
 ; OTHER INFORMATION: protein kinase
 US-08-557-006C-37

Query Match 7.88; Score 182.6; DB 4; Length 1736;
 Best Local Similarity 53.38; Pred. No. 7.7e-33;
 Matches 408; Conservative 0; Mismatches 354; Indels 3; Gaps 1;

Oy 92 TGGGCCCTATCGGCTGAGAGAGAGCTGGGCAAGAGACAGAGGGCTGGTTAACTCG 151
 Db 38 TCGGACACTAGCTGCTGGGCGACACGCTGGGCGTGGGCACTTCGGCAAGAGAAATGG 97
 Oy 152 GGGTCACCTGCATCAGCGGCTGAGAGAGTCCCATCAAGATCGTGAACCGGAGAACCTGT 211
 Db 98 GAGAACATCAATTAACAGGCCATTAAGTGGCAGTTAAATCTTAATATAGACAGAAATTC 157
 Oy 212 CGAGT---CGGTGCTGATGAAGGTGAGCGGAGATGCCATCCCTGAAGCTCATCGAAC 268
 Db 158 GCACTTATAGATGTTGTTGGAATAAATAAAGCAAAATTCMAAATCTTAAACCTCTTCGTC 217
 Oy 269 ACCGACATGCTCTCAAGCTCCACGAGCTCTACGAGAACAAATATTTGTACCTGGTTC 328
 Db 218 ATCTCTATATATCAAACTATACAGGATGATGACATCCCAACAGATTTTATATGTA 277
 Oy 329 TGAAGCAGCTCGGGGGGTAGCTATTGCACTACCTGGTAAAGAGGGAGACTGACGC 388
 Db 278 TGAATATGTCTGAGAGTGAATATTTGACTACATCTGTAGCATGAGAGGGGTGAAG 337
 Oy 389 CCAAGGAGCCCGAAAGTTCTCCGCCAGATTTGTCTGCGCTGAGACTCTGCGACAGCT 448
 Db 338 AGATGAGAGCCGAGCGGCTCTTTCAGCAGATTCCTGCTGTGATTTACTGTCAATAGGC 397
 Oy 449 ACTCATCTGCGCAGAGAGACTTAAGCCGAGAACCTGTTTGGATGAGAAACAACA 508
 Db 398 ATATGGTTGTTCATCAGACCTGAAACAGAGAAATGTCTGTGGTTCACACATATAAG 457
 Oy 509 TCCGCAATTCAGACTTGGCATGCGGCTCCCTGACAGTGGGGAGACCCCTCTGAGACCA 568
 Db 458 CCAAGATAGCCGATTTTGGATTATCTAATATGATGTCAAGATGTAATTTCTGAGAACTA 517

Oy 569 GCTGGGGTCCCGCCCATATGCTGTCCAGAGTGAATTAAGGGGAAATATGATGCC 628
 Db 518 GTTCCGATCTCAAAATTAATGACACGCTGAATCTCATCTGATGAGAGTTC 577
 Oy 629 GCCGGCAGACATGTGAGAGCTGTGGAGCATCCTTCGCGCCCTGCTCGGGGGCTGCG 688
 Db 578 CTGAAGTTGATATCTGAGCTGTGTGTATCTTATGCTCTTTTGTGGCACCTCC 637
 Oy 689 CTTTGTATGACAGACCTCCGCGAGCTGTGAGAGGTGAAGCGGCTCTTCACACA 748
 Db 638 CATTTATGATGAGACATGTACCTAGCTATATTAAGAAGATCCGAGGGGTGCTTTTANA 697
 Oy 749 TGCCCCACTTATCTCTCAATTTCCAGAGCTCTCTGAGGGAATGATCGAATGAGC 808
 Db 698 TCCGAAATATCTCAATGCTGTCTGCGCACTCTCTGATGATATGCTGACAGGTGACC 757
 Oy 809 CCGAAGAAAGGCTCAGCTTGAGCAAAATTCAGAAACATCCTTGGT 853
 Db 758 CACTGAAGCGAGCAACTATCAAGACATTAAGAGACATGATGAT 802

RESULT 9

US-08-557-006C-36
 ; Sequence 36, Application US/08557006C
 ; Patent No. 6258547
 ; GENERAL INFORMATION:
 ; APPLICANT: Berl, Rajindar K.
 ; APPLICANT: Carling, David
 ; APPLICANT: Forster, Robert A.
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
 ; FILE REFERENCE: NGAP/PH37588/OST
 ; CURRENT APPLICATION NUMBER: US/08/557,006C
 ; CURRENT FILING DATE: 1996-03-06
 ; PRIOR APPLICATION NUMBER: PCT/GB94/01093
 ; PRIOR FILING DATE: 1994-05-20
 ; PRIOR APPLICATION NUMBER: GB 9310489.1
 ; PRIOR FILING DATE: 1993-05-21
 ; PRIOR APPLICATION NUMBER: GB 9318010.7
 ; PRIOR FILING DATE: 1993-08-31
 ; NUMBER OF SEQ ID NOS: 44
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 36
 ; LENGTH: 1783
 ; TYPE: DNA
 ; ORGANISM: Human AMP protein kinase
 US-08-557-006C-36

Query Match 7.88; Score 182.6; DB 4; Length 1783;
 Best Local Similarity 53.38; Pred. No. 7.8e-33;
 Matches 408; Conservative 0; Mismatches 354; Indels 3; Gaps 1;

Oy 92 TGGGCCCTATCGGCTGAGAGAGAGCTGGGCAAGAGACAGAGGGCTGGTTAACTCG 151
 Db 39 TCGGACACTAGCTGCTGGGCGACACGCTGGGCGTGGGCACTTCGGCAAGAGAAATGG 98
 Oy 152 GGGTCACCTGCATCAGGGGTGAGAGGTCCCATCAAGATCGTGAACCGGAGAACCTGT 211
 Db 99 GAGAACATCAATTAACAGGCCATTAAGTGGCAGTTAAATCTTAATATAGACAGAAATTC 158
 Oy 212 CGAGT---CGGTGCTGATGAAGGTGAGCGGAGATGCGCATCTCTGAAGCTCATCGAAC 268
 Db 159 GCACTTATAGATGTTGTTGGAATAAATAAAGCAAAATTCMAAATCTTAAACCTCTTCGTC 218
 Oy 269 ACCGACATGCTCTCAAGCTCCACGAGCTCTACGAGAACAAATATTTGTACCTGGTTC 328
 Db 219 ATCTCTATATATCAAACTATACAGGATGATGACATCCCAACAGATTTTATATGTA 278
 Oy 329 TGAAGCAGCTCGGGGGGTAGCTATTGCACTACCTGGTAAAGAGGGAGACTGACGC 388
 Db 279 TGAATATGTCTGAGAGTGAATATTTGACTACATCTGTAAAGCATGAGACGGGTGAAG 338
 Oy 389 CCAAGGAGCCCGAAAGTTCTCCGCCAGATTTGTCTGCGCTGAGACTCTGCGACAGCT 448

```

Db 339 AGATGAGACCGGCGCTTTTCAGCAGATTCTGCTGCTGATTAAGTCTATAGCC 398
QY 449 ACTCCATCTGCGACAGAGACCTAAAGCCGAGAACCTGTTTGGATGAGAAAACACA 508
Db 399 ATATAGTTGTTTCACTGAGACCTGAAACAGAAATGCTCTGTTGATGACATGAAATG 458
QY 509 TCCGATTCGACAGATGCGCATGCGGCTCCCTGCAGGTGGGGACACCTCTCGAGACCA 568
Db 459 CCAAGATAGCCGATTTGGATTATCTATATGATGTCAGATGGTGAATTTCTGAGAACTA 518
QY 569 GCTGCGGCGCCCGCCATTATGCTGTCAGAGGTGATTAAGGGGAAAATATGATGGCC 628
Db 519 GTTCCGATCTCCAAATTAATGACAGACCTGAAATCATCTCAGCCAGATTTATGACAGTTC 578
QY 629 GCCGGGACAGATGAGAGCTGTGAGTCACTCTTCCGCTGCTGCGGGGCTCTGTC 688
Db 579 CTGAAGTTGATCTGAGAGCTGTGAGTCACTCTTCCGCTGCTGCGGGGCTCTGTC 638
QY 689 CCTTGTATGACAGACCTCTCCGACCTGCTGAGAGAGTGAACGGGCGCTCTTCCACA 748
Db 639 CATTGTATGATGACAGATGATCTACCTGATTTTAAGAAATCCGAGGGGCTCTTTTATA 698
QY 749 TGGCCCATCTCATCTCCAGATTCGACAGAGCTCTGAGGGGAAATGAATGCAAGTGGAGC 808
Db 699 TCCAGATATCTCAATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 758
QY 809 CCGAAAAAGGCTCAGTCTGAGCAAAATTCAGAAACATCCTTGTGT 853
Db 759 CACTGAACGAGCAACTATCAAAAGCAATGAAGAGCATGAATGTGT 803

```

RESULT 10
US-08-677-298-1
; Sequence 1, Application US/08677298
; Patent No. 5863729

GENERAL INFORMATION:

APPLICANT: Pimica-Worms, Helen
TITLE OF INVENTION: DNA SEQUENCES ENCODING HUMAN TCAR-1
NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESSES:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201

CITY: Boulder
STATE: CO
COUNTRY: USA

ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/677,298
FILING DATE: 09-JUL-1996

CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:

NAME: Caruthers, Jennie M.
REGISTRATION NUMBER: 34,464
REFERENCE/DOCKET NUMBER: 9-96

TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 2698 base pairs
TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: not relevant

MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO

ANTI-SENSE: NO
FEATURE:

NAME/KEY: CDS
LOCATION: 376..2565
US-08-677-298-1

Query Match
Best Local Similarity 7.6%; Score 177.8; DB 2; Length 2698;
Pred. No. 1.1e-31;
Matches 404; Conservative 0; Mismatches 377; Indels 0; Gaps 0;

```

QY 92 TGGGCCCCCTATCGCTGAGAGAGAGCTGGGCAAGAGACAGACAGGCTGGTTAACTCG 151
Db 533 TCGGAACACTACAGACTTGTGAAAACATCGGCAAGGGAAATTTTGCAAAATTAATGG 592
QY 152 GGGTCCAGCTCAGTACCGGGTACAGAGTCCGCAATCAAGTGTGAACCGGGAGAGCTGT 211
Db 593 CAAGACTATATCTTACAGGAGAGAGAGGTTGCAATTAATTAATGACAAACTAGTTGA 652
QY 212 CGAGAGTGGTGTGATGAAGTGAAGGGGAGATCCGCAATCTGTAAGTCTATGAGACCC 271
Db 653 ATCCAAAGCTTACAAAGCTCTTACAGAGAGTGAAGTGAATATATGAATTTAAATCATTC 712
QY 272 CACATGCTCTCAAGCTCCAGAGCTTACGAGACAGAAATATTTGACCTGGTTCTGG 331
Db 713 CCAATATATGAAAGTTATTTGAAGTCAATGAAGTGAAGAAACACTTCACTATATG 772
QY 332 AGCAGCTCTCGGGGGGTGAGCTATTTGACTACTGTAAGAGAGGGAGACTGACGCCCA 391
Db 773 AATATGCAAGTGAAGAGGAGATATTTGATTTGTTGTCACATGGCAGATGAAGAAA 832
QY 392 AGGAGCGCGGAAAGTTCTCCGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 451
Db 833 AAGAGAGAGATCTAAATTTAGACAGATGCTGTGCAATCAATCACTCAGAAAC 892
QY 452 CCATCTGCGACAGAGACCTAAAGCCGAGAACTCTTTTGGATGAGAAAACATATCC 511
Db 893 GGAATCGATACAGAGACTGAAAGCTGAATAATATTTGATGATGCCATATGAACATTA 952
QY 512 GCATTCGAGACTTGGGATGAGGCTCCCTGCAAGTGGGGGAGAGGCTCTGAGACCACT 571
Db 953 AATATGAGATTTGCTGTTTGAAGATGAATTTACTGTTGGGGTGAATCTGACACAGTTT 1012
QY 572 GGGGTCGCCCATTTATGCGTGTCCAGAGTGAATTAAGGGGAAATATATATGCGGCC 631
Db 1013 GTGGCAGTCTCAATAGCAGACACTGAGCTCTTCCAGGGGAAAGAAATATACGGCCAG 1072
QY 632 GGGCAGCATGTGAGAGTGTGAGATCTCTTCCGCTGCTGCTGCTGCTGCTGCTGCTG 691
Db 1073 AAGTGAATGTGTGAGAGTGTGAGATCTCTTATACACTGATGATGCTGCTGCTGCT 1132
QY 692 TTGATGACGAGAACTTCCGCGAGCTGTGAGAGAGTGAAGCGGGGCTTCCACATGC 751
Db 1133 TTGATGAGGCAAAACCTAAAGAACTGAGAGAGATTAAGAGGGAATTAAGAAATTC 1192
QY 752 CCGACTTCAATCTCCAGATTTGCCAGAGCTCTGAGGGGAAATATGAATGAGACCG 811
Db 1193 CTTTCAATGTCTACAGAGTGTGAAGAACTTCAAAAGCTTCTGCTGCTGAATATCAA 1252
QY 812 AAAAAGGCTCAGTCTGAGCAAAATTCAGAAACATCTCTGTACTTACAGCGGAAACG 871
Db 1253 TTAAGCGGCACTCTAGAGCAAAATCATGAAGAGCAGAGTGAATCAATGACAGGGCATG 1312
QY 872 A 872
Db 1313 A 1313

```

RESULT 11

US-08-557-006C-44
; Sequence 44, Application US/08557006C
; Patent No. 6258547

GENERAL INFORMATION:

APPLICANT: Berl, Rajindar K.
APPLICANT: Carling, David
APPLICANT: Forster, Robert A.

Query Match 5.8%; Score 135.6; DB 2; Length 2132;
 Best Local Similarity 54.1%; Pred. No. 4,7e-22;
 Matches 360; Conservative 0; Mismatches 279; Indels 27; Gaps 3;

```

QY 78 GCACGCCCAATATGTGGGCCCTATCGGCTGGAAGAAGACGCTGGGCAAGAGCAACAGCAGG 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108 GCAGAGAGAGCTGGAGACCATTTATGAGATGGGGAGAGCTGGGACACGGCCAGTTTGC 167
QY 138 GCTGTAAACTCGGGGTCCACATCAGCGGGTCAAGAGTCCGCATCAAGATGCTGAA 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 168 GATCTGCGGAAGTCCGGGCAAGAGGACAGGAGTACGACGCAAGTTCATCAA 227
QY 198 CCGGAGAAAGCTGTC-----GAGTCGGTCTGATGAAGTGAAGTGAAGGAGAT 245
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 228 GAAGCCCGCTGTATCATCCAGCCGCGTGGGGTGAAGCCGGAGAGATCGAGCGGAGGT 287
QY 246 CGCATCTCTGAAGCTATCGAACACCCACATGTCTCAAGCTCCAGCAGCTTACGAGAA 305
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 288 GAACATCTCTCGGAGATCCGGACCCCAACATCATCCTTCGACGACATCTTCGAGAA 347
QY 306 CAAGAAATATTTGACTGTGTTGAGACAGCTCTCGGGGGGTGAGCTATTGCACTACT 365
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 348 CAAGAGGAGAGCTGTCTCTCATCTGAGAGCTGTCTCTCGGGGAGCTCTTTGACTTCT 407
QY 366 GGTAAAGAGGGGAGACTGACGCCCAAGAGAGGCCGAAAGTTCTTCCGACAGATTGTGC 425
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 408 GCGGAGAAAGAGTCTGCTGAGAGAGAGAGAGAGCCAGCTTCTCAAGAGATCTGGA 467
QY 426 TGGGCTGAGCTTGTGACACAGCTACTCCATCTGACACAGAACCTTAAGCCGGAACCT 485
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 468 CGCGCTCACTACTGCTGCTTAAGCGATCGACACTTTTACGAGAGCCGGAAGAACT 527
QY 486 GCTTTGGATGAGAAAAAC-----ACATCCGCAATTGAGAGACTTGGGATGGC 533
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 528 CATGCTGCTGAGAACAGAGCCCAACCCAGATCAAGTCACTTCACTTGGATGCG 587
QY 534 GTCCCTGACAGTGGGGAGACGCTCTGAGAGACAGCTGCGGGTCCGCCATTATGCGTG 593
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 588 GCACAAGATCGAGCGGGGAGAGAGTTCAAGAACATCTTCGGACACCCGAGTTGTGGC 647
QY 594 TCCAGAGGTATTAAGGGGAAAAAATATGATGGCCGCGGGGACAGATGAGAGCTGTGG 653
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 648 CCCAGAGATTGTGAATCATATGAGCCGCTG---GGCTTGAGAGCGGACATGTGAGCATCGG 704
QY 654 AGTCATCTCTTCCCTGCTGCTGTGGGGGCTCTGCTCTTTGATGAGAGCAACCTCCGCCA 713
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 705 TGTCACTACCTATATCTCTCTGAGCGGTGATCCCGTTCTCTGGGCGAGACCAAGCAGGA 764
QY 714 GCTGCT 719
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 765 GACGCT 770

```

RESULT 14
 US-09-186-277-3

```

; Sequence 3, Application US/09186277
; Patent No. 6171841
; GENERAL INFORMATION:
; APPLICANT: AKIRA, SHIZUO
; APPLICANT: KAWAI, TARO
; TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
; FILE REFERENCE: 081356/0128
; CURRENT APPLICATION NUMBER: US/09/186, 277
; CURRENT FILING DATE: 1998-11-05
; EARLIER APPLICATION NUMBER: JP97/261589
; EARLIER FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2132
; TYPE: DNA
; ORGANISM: Homo sapiens

```

FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (94)..(1455)
 US-09-186-277-3

Query Match 5.8%; Score 135.6; DB 4; Length 2132;
 Best Local Similarity 54.1%; Pred. No. 4,7e-22;
 Matches 360; Conservative 0; Mismatches 279; Indels 27; Gaps 3;

```

QY 78 GCACGCCCAATATGTGGGCCCTATCGGCTGGAAGAAGACGCTGGGCAAGAGCAACAGCAGG 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108 GCAGAGAGAGCTGGAGACCATTTATGAGATGGGGAGAGCTGGGACACGGCCAGTTTGC 167
QY 138 GCTGTAAACTCGGGGTCCACATCAGCGGGTCAAGAGTCCGCATCAAGATGCTGAA 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 168 GATCTGCGGAAGTCCGGGCAAGAGGACAGGAGTACGACGCAAGTTCATCAA 227
QY 198 CCGGAGAAAGCTGTC-----GAGTCGGTCTGATGAAGTGAAGTGAAGGAGAT 245
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 228 GAAGCCCGCTGTATCATCCAGCCGCGTGGGGTGAAGCCGGAGAGATCGAGCGGAGGT 287
QY 246 CGCATCTCTGAAGCTATCGAACACCCACATGTCTCAAGCTCCAGCAGCTTACGAGAA 305
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 288 GAACATCTCTCGGAGATCCGGACCCCAACATCATCCTTCGACGACATCTTCGAGAA 347
QY 306 CAAGAAATATTTGACTGTGTTGAGACAGCTCTCGGGGGGTGAGCTATTGCACTACT 365
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 348 CAAGAGGAGAGCTGTCTCTCATCTGAGAGCTGTCTCTCGGGGAGCTCTTTGACTTCT 407
QY 366 GGTAAAGAGGGGAGACTGACGCCCAAGAGAGGCCGAAAGTTCTTCCGACAGATTGTGC 425
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 408 GCGGAGAAAGAGTCTGCTGAGAGAGAGAGAGCCAGCTTCTCAAGAGATCTGGA 467
QY 426 TGGGCTGAGCTTGTGACACAGCTACTCCATCTGACACAGAACCTTAAGCCGGAACCT 485
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 468 CGCGCTCACTACTGCTGCTTAAGCGATCGACACTTTTACCTGAGAGCCGGAAGAACT 527
QY 486 GCTTTGGATGAGAAAAAC-----ACATCCGCAATTGAGAGACTTGGGATGGC 533
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 528 CATGCTGCTGAGAACAGAGCCCAACCCAGATCAAGTCACTTCACTTGGATGCG 587
QY 534 GTCCCTGACAGTGGGGAGACGCTCTGAGAGACAGCTGCGGGTCCGCCATTATGCGTG 593
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 588 GCACAAGATCGAGCGGGGAGAGAGTTCAAGAACATCTTCGGACACCCGAGTTGTGGC 647
QY 594 TCCAGAGGTATTAAGGGGAAAAAATATGATGGCCGCGGGGACAGATGAGAGCTGTGG 653
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 648 CCCAGAGATTGTGAATCATATGAGCCGCTG---GGCTTGAGAGCGGACATGTGAGCATCGG 704
QY 654 AGTCATCTCTTCCCTGCTGCTGTGGGGGCTCTGCTCTTTGATGAGAGCAACCTCCGCCA 713
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 705 TGTCACTACCTATATCTCTCTGAGCGGTGATCCCGTTCTCTGGGCGAGACCAAGCAGGA 764
QY 714 GCTGCT 719
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 765 GACGCT 770

```

RESULT 15
 US-09-159-385-4

```

; Sequence 4, Application US/09159385
; Patent No. 5958748
; GENERAL INFORMATION:
; APPLICANT: AKIRA, SHIZUO
; APPLICANT: KAWAI, TARO
; TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
; FILE REFERENCE: PH-569
; CURRENT APPLICATION NUMBER: US/09/159, 385
; CURRENT FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: JP97/261589
; EARLIER FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0

```

```

: SEQ ID NO 4
:
: LENGTH: 1429
:
: TYPE: DNA
:
: ORGANISM: Mus musculus
:
: FEATURE:
:
: NAME/KEY: CDS
:
: LOCATION: (10)..(1353)
:
: OS-09-159-385-4

```

Query Match	5.6%	Score 130.6;	DB 2;	Length 1429;
Best Local Similarity	56.5%	Pred. No. 5.8e-21;		
Matches 291; Conservative	0;	Mismatches 209;	Indels 15;	Gaps 2

QY	223	CTGATGAAGGTGGAGACCGGAAATGGCATCTCTGAAAGTCAATCGCAACCCCAATGTCTC	282
Db	181	CGGGAGGAGATCGAAACCGGAGGTGAGATCTCTGGCCAGATCCGCCAACCCCAATCATATA	240
QY	283	AAGTCCACGACGTCTACGAGAACAGAAATATTGTACTGGTTCTTGAGACGTCGC	342
Db	241	ACACTGCTATGACGTTCCTCGAAMACACAGATGTGGTCTGATCTCTGGAGCTGTCTCC	300
QY	343	GGGGGTGAGCTATTCGACTCTCTGTAAAGAGGGGAGACTGAGAGGCCCAAGAGGGCCCGA	402
Db	301	GGTGGCGAGCTTTTTCGACTTTCCTGTGCGCCGAGAGAGATCATTTGAGGAGATGAGGCCAG	360
QY	403	AAGTTCTTCGCGCAGATTGTGTCTGCGCTGAGACTTCTGCCACAGCTTACTCATCTGCGAC	462
Db	361	CAGTTCCCAACAACAAATCCATGAGCGGTGCTACACTCTGCACCTCGAACGCGCATCGCAC	420
QY	463	AGAGACCTAAAGCCCGAGAACCTCTTTTGGATGAGAAAAACAAATCTCCG-----	513
Db	421	TTTGACTCGAAGCCCGAAMCATATATGTGTGGACAAAGCACGACGACACCCCGCATTT	480
QY	514	---ATTGCAGACTTCGGGATGCGCTCCCTCGGAGGTGGGGGACACGCTCTCGAGACCAGC	570
Db	481	AAGCTCATCGACTTTGGCATGCGCCACAGATCGAAGGCTGGCACCGAGTTCAAGAACATC	540
QY	571	TGGGGGTCCCCCATTTATGCGTGTCCAGAGGTATTAAAGGGGAGAAAAATATGATGGCCGC	630
Db	541	TTTGGCACACCCGAGTTGTGCGCCCCCGAGATCGTGAACATATGAGCCA---CTTGGCTTG	597
QY	631	CGGGCAGACATGTGGAGCTGTGAGTCACTCTCTTCGCCCTCTGCTGCTGGGGGCTCTTCCC	690
Db	598	GAGGCTGACATGTGGAGCATTTGGGCTCATCTACATATCTCTCTGACGGAGGGTCCCA	657
QY	691	TTTGTATGACGAACATCTCCGCCCACTCTGTGAGAA	725
Db	658	TTTCTGGGCGAGACCAAGCAGGAACGCTGACGA	692

Search completed: April 22, 2003, 00:26:08
Job time : 66.0237 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 20:07:11 ; Search time 2126.25 Seconds

(without alignments)
17777.901 Million cell updates/sec

Title: US-10-003-690-3

Perfect score: 2334
1 atggtgtccgggccaaga.....ccaagggaccctctgcc 2334

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 33308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlun:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gD_estl:*
10: gD_estc2:*
11: gD_estc3:*
12: gD_estc4:*
13: gD_estc5:*
14: gD_estc6:*
15: em_estfun:*
16: em_estlom:*
17: gD_gss:*
18: em_gss_hum:*
19: em_gss_liv:*
20: em_gss_pln:*
21: em_gss_vtc:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	759.8	32.6	889	14	B0434571 AGENCOURT
2	723.6	31.0	881	14	B0717903 AGENCOURT
3	651.2	27.9	831	12	BF529743 AGENCOURT
4	602.2	25.8	658	12	BG819795 AGENCOURT
5	592.4	25.4	698	14	BM948524 UT-M-EGOP
6	574.8	24.6	742	14	BM944329 UT-M-EGOP

7	543	23.3	671	14	B0178547	B0178547 UT-M-EVO-
8	540.4	22.6	658	14	B0573641	B0573641 UT-M-EDO-
9	527.6	22.2	836	9	AL538014	AL538014 AL538014
10	522.4	22.4	936	12	BF529630	BF529630 602043855
11	510	21.9	510	14	B0086295	B0086295 1120401.Y
12	490.8	21.0	898	12	BF529174	BF529174 602042051
13	490.2	21.0	600	14	B0444032	B0444032 UT-M-EVO-
14	487.8	20.9	753	12	BF796285	BF796285 602258590
15	476.2	20.4	614	14	BM944236	BM944236 UT-M-EGOP
16	460.4	19.7	728	14	B0769984	B0769984 UT-M-FTO-
17	444	19.0	492	12	BF688432	BF688432 IL3-ET011
18	440.4	18.9	597	13	BF689059	BF689059 602295163
19	439	18.8	453	12	BF364892	BF364892 QVO-NH114
20	429.4	18.4	733	14	B0178695	B0178695 UT-M-EVO-
21	426.6	18.3	508	14	BM950539	BM950539 UT-M-EGOP
22	415.4	17.8	521	14	B0442731	B0442731 UT-M-EVO-
23	406.4	17.4	693	14	B0443644	B0443644 UT-M-EVO-
24	406	17.3	961	14	W41136	W41136 MC38F03.r1
25	404.4	17.3	551	12	BG715357	BG715357 602677384
26	374.4	16.0	488	14	W60403	W60403 602677384
27	366.2	15.7	793	9	AL135353	AL135353 GHI3047.5
28	365	15.6	596	10	BE222941	BE222941 hu44907.x
29	352.2	15.1	427	14	W18584	W18584 mb88c11.r1
30	341.8	14.6	418	10	AM122891	AM122891 UT-M-BH2
31	338	14.5	712	14	B0180079	B0180079 UT-M-EVO-
32	336.2	14.4	925	14	B0435328	B0435328 AGENCOURT
33	331.8	14.2	609	14	BM949588	BM949588 UT-M-EGOP
34	331.2	14.0	344	12	BF931102	BF931102 CML-NR004
35	326	14.0	367	12	BE763502	BE763502 OVI-TM004
36	325	13.9	337	13	BM129669	BM129669 1121e10.Y
37	319.8	13.7	394	14	W75642	W75642 me49a09.r1
38	318.2	13.6	1177	13	BM466629	BM466629 AGENCOURT
39	307.2	13.2	733	12	BG395625	BG395625 602458251
40	307.2	13.2	1882	11	BC017182	BC017182 Homo sap1
41	306	13.1	546	14	BM944658	BM944658 UT-M-EGOP
42	305.2	13.1	563	10	BE251924	BE251924 601107590
43	302.6	13.0	334	12	BF371253	BF371253 RC6-FW013
44	293.2	12.6	740	12	BF222739	BF222739 7q32f02.x
45	288.4	12.4	293	9	AA326737	AA326737 EST29963

ALIGNMENTS

RESULT 1
B0434571
LOCUS
DEFINITION
AGENCOURT_7917755 NIH_MGC_67 Homo sapiens CDNA clone IMAGE:6154749
5', mRNA sequence.
889 bp mRNA linear EST 24-MAY-2002
B0434571 GI:21173647

ACCESSION
B0434571
VERSION
B0434571.1
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 889)
AUTHORS
NIH-MGC <http://mgi.cni.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1AM13496 row: e column: 22
High quality sequence stop: 681.
Location/Qualifiers
1..889

FEATURES
source

BASE COUNT	207	a	246	c	274	g	160	t	2 others
ORIGIN	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:6154749" /clone_1bp="NH.MGC.67" /issue_type="retinoblastoma" /lab_host="DH10B (phage-resistant)" /note="Organ: eye; Vector: PCWV-SPO86; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."								

Query Match	32.6%	Score 759.8	DB 14	Length 889
Best Local Similarity	98.5%	Pred. No. 2.7e-154		
Matches 788	Conservative	0	Mismatches 8	Indels 4
				Gaps 2

OY	104	GGGGGAAAGACGCTGGGGCAAGAGACAGGGCGGTAAACATCGGGGTCCACTGGCA	163
Db	9	GGGTGGAAAGAACCCCTGGGCAAGAGACAGACGGGCTGGTTAAACTCGGGGTCCACTGGCA	68
OY	164	TCACGGGTCAAGAAAGTGGCCATCAAGATCGTGAACCGGGAGAAAGCTGTCCGAATCGGTGC	223
Db	69	TCACGGGTCAAGAAAGTGGCCCATCAAGATCGTGAACCGGGAGAAAGCTGTGGATCGGTGC	128
OY	224	TGATGAAGTGGAGCCGGGAGATTCGCCATCCTGAAGCTATCGAACAACCCATGTCTCTCA	283
Db	129	TGAATGAAGTGGAGCCGGGAGATTCGCCATCCTGAAGCTATCGAACAACCCATGTCTCTCA	188
OY	284	AGCTCCACAGAGTGTACGAGAACAAAGAAATATTTGTACTCGTGTCTGGACACGTCTCGG	343
Db	189	AGCTCCACAGAGTGTACGAGAACAAAGAAATATTTGTACTCGTGTCTGGACACGTCTCGG	248
OY	344	GGGGTGAAGCTATTTCGACTACTGGTAAAGAAAGGGAGACTGACGCCAGAGGCCCGAA	403
Db	249	GGGGTGAAGCTATTTCGACTACTGGTAAAGAAAGGGAGACTGACGCCAGAGGCCCGAA	308
OY	404	AGTTCTTCCGCGACAATGTGTCTCGGTGGACGTTCTCCACAGTACTCCATGTGGCACA	463
Db	309	AGTTCTTCCGCGACAATGTGTCTCGGTGGACGTTCTCCACAGTACTCCATGTGGCACA	368
OY	464	GAGACCTAAAGCCCGAGAACCTCGTTTTGGATGAGAAAAACAACATCCGATTCGACACT	523
Db	369	GAGACCTAAAGCCCGAGAACCTCGTTTTGGATGAGAAAAACAACATCCGATTCGACACT	428
OY	524	TCGGCATATGGCGTCCCTGCAAGTGGGGGACACCTCTCTGGAGACACGCTGGGGTCCCCC	583
Db	429	TCGGCATATGGCGTCCCTGCAAGTGGGGGACACCTCTCTGGAGACACGCTGGGGTCCCCC	488
OY	584	ATTATGCTGTCCAGAGGTGATTTAAGGGGAAAAATTGATGGCGCGCGGACAGACTGT	643
Db	489	ATTATGCTGTCCAGAGGTGATTTAAGGGGAAAAATTGATGGCGCGCGGACAGACTGT	548
OY	644	GGAAGTGTGAGTCATCCYTCGCTCGCTCGTGGGGGCTCTGACCCTTTTGATGACGACA	703
Db	549	GGAAGTGTGAGTCATCCYTCGCTCGCTCGTGGGGGCTCTGACCCTTTTGATGACGACA	608
OY	704	ACCTCCGCGACACTCTGGAAGAAAGGTGAAAGGGGCGCTCCACATGCCCCCATTCATTTC	763
Db	609	ACCTCCGCGACACTCTGGAAGAAAGGTGAAAGGGGCGCTCTTCCACATGCCCCCATTCATTTC	668
OY	764	CTCCAGATTGGCAAGGCTCTCTGAGGGGAATGATCGAATGGAGCCCGAAAAAAGGCTCA	823
Db	669	CTCCAGATTGGCAAGGCTCTCTGAGGGGAATGATCGAATGGAGCCCGAAAAAAGGCTCA	728
OY	824	GTCCTGAGACAAATTCAGAAACATCTTTGG--TACTGAGCGGGGAAAAACAGACCCAGACC	880
Db	729	GTCCTGAGACAAATTCAGAAACATCTTTGGTGTACNCTAGGGCGGGGAAAAACAGACCCAGACC	788
OY	881	CGTG--CGTGGAGCCAGCCCC 899	
Db	789	CGTGCCCTGGAGCCAGCCCC 808	

RESULT 2	881 bp	mRNA	linear	EST 16-JUL-2007
LOCUS	BO7117903			
DEFINITION	AGNCDCDRT.8241190 Lupsk1-sympathetic_c.trunk Homo sapiens cDNA clone			
IMAGE:	6186946 5', mRNA sequence.			
ACCESSION	BO7117903			
VERSION	BO7117903.1	GI:21856800		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			

REFERENCE	1 (bases 1 to 881)
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Department of Biology National Institutes of Health Bethesda, MD 20892-0445 USA Tel: 301-495-4600 Fax: 301-495-4600 E-mail: strausberg@ncl.nih.gov

Tissue Procurement: Dr. James R. Lipsky
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LNA13580 row: c column: 11
High quality sequence: 620.

FEATURES	Location/Qualifiers
source	1. .881

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6186946"
/clone_lib="lupsk1-sympathetic_trunk"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/notes="Vector: PCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACGACGACGCGCCG-3' and
5'-GACATGCTTCAGATCGGACGCGCCGCT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."
```

BASE COUNT	192 a	260 c	271 g	155 t	3 others
ORIGIN					

Query Match	31.0%	Score 723.6	DB 14	Length 881
Best Local Similarity	95.3%	Pred. No. 1.9e-146		
Matches 755	Conservative 0	Mismatches 36	Indels 1	Gaps 1

QY	485	TGCTTTTGATGAGAAAAACAATCCGATTTGGAGACTTGGATGGCGTCCCTGGAGG	544
Db	1	TGCTTTTGATGAGAAAAACAATCCGATTTGGAGACTTGGATGGCGTCCCTGGAGG	60
QY	545	TGGGGGACAGCCTCCTGTGAGACCAAGCTGCGGGTCCCCCATTTATGCGTGTCCAGAGGTGA	604
Db	61	TGGGGGACAGCCTCCTGTGAGACCAAGCTGCGGGTCCCCCATTTATGCGTGTCCAGAGGTGA	120
QY	605	TTTAAGGGGAAAAATATATGATGGCGCGCGGAGACATGTGGAGCTGTGAGATTCATCTCT	664
Db	121	TTTAAGGGGAAAAATATATGATGGCGCGCGGAGACATGTGGAGCTGTGAGATTCATCTCT	180
QY	665	TGCGCCCTGCTGTGGGGGCTCTGCCCTTTATATGACGACAACCTCCGACAGTCTGCTGGAGA	724
Db	181	TGCGCCCTGCTGTGGGGGCTCTGCCCTTTATATGACGACAACCTCCGACAGTCTGCTGGAGA	240
QY	725	AGGTGAACCGGGGGCTCTTCACATGGCCCACTTCATTCCTCCAGATTTGCCAGAGCCCTCC	784


```

Db 241 AGGTGAACGGGGGCTTCCACATGCCCCCACTTCATCTCCTCCAGATTGCCAGACCTCC 300
QY 785 TGAGGGGAATGATGAGAGTGGAGCCCGAAAAAGGCTCAGTCTGGAGCAAAATTCAGAAAC 844
Db 301 TGAGGGGAATGATGAGAGTGGAGCCCGAAAAAGGCTCAGTCTGGAGCAAAATTCAGAAAC 360
QY 845 ATCTGTGCTACCTGAGGGGGAACACAGACGCCAGACCCCTGCTGAGACGACCCCTGGGC 904
Db 361 ATCTGTGCTACCTGAGGGGGAACACAGACGCCAGACCCCTGCTGAGACGACCCCTGGGC 420
QY 905 GCCGGGAGCCATGCGAGGCTGCTCCATCCACAGAGAGCTGGACCCCGACGCTAGAGA 964
Db 421 GCCGGGAGCCATGCGAGGCTGCTCCATCCACAGAGAGCTGGACCCCGACGCTAGAGA 480
QY 965 GCATGCGATCAGTGGGCTGCTTACAGGACCGCAGAGAGGCTCATGCGAGCTGCGCAGTG 1024
Db 481 GCATGCGATCAGTGGGCTGCTTACAGGACCGCAGAGAGGCTCATGCGAGCTGCGCAGTG 540
QY 1025 AGGAGGGAACCAAGAAAGATGATATATATATATATATATATATATATATATATATATAT 1084
Db 541 AGGAGGGAACCAAGAAAGATGATATATATATATATATATATATATATATATATATATAT 600
QY 1085 CCAAGCTGAGAGACCAAGACCTGCTCCCGCGAAATGATGTTGACCCCGCCCGG-AAAGCT 1143
Db 601 CCAAGCTGAGAGACCAAGACCTGCTCCCGCGAAATGATGTTGACCCCGCCCGG-AAAGCT 660
QY 1144 GTGAGATCTCCCATGCTGAGAGCGCTGACGGAAGCGCGACCAAGAGCAAGTCCATGAA 1203
Db 661 GTGAGATCTCCCATGCTGAGAGCGCTGACGGAAGCGCGACCAAGAGCAAGTCCATGAA 720
QY 1204 GTCTGTGCAATCCCAATGCTGCGGGGGGTGTGGCTCCCTGTACCCACCCGAGGGGCTTG 1263
Db 721 GTCTGTGCAATCCCAATGCTGCGGGGGGTGTGGCTCCCTGTACCCACCCGAGGGG 780
QY 1264 GAGATGGCCAG 1275
Db 781 GCCTTGCCAAAG 792

RESULT 3
LOCUS BF529743 831 bp mRNA linear EST 11-DEC-2000
DEFINITION 602044002F1 NCL_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4181688
5', mRNA sequence.
ACCESSION BF529743 GI:11617106
VERSION BF529743
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Chordata; Craniata; Vertebrata; Euteleostomi;
            Eukaryota; Metazoa; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 831)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
            Email: cgapds@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L14M9495 row: c column: 01
            High quality sequence stop: 736.
            Location/Qualifiers
                1. 831
FEATURES
    source
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:4181688"
        /clone_11b="NCL_CGAP_Brn67"
        /tissue_type="anaplastic oligodendroglioma with 1p/19q

```

```

loss"
/lab_host="DH10B (T1 phage-resistant)"
/Note="Organ: brain; Vector: pCMV-SpRTE; site_1: NotI;
site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCL_CGAP Library."
BASE COUNT 185 a 237 c 258 g 151 t
ORIGIN
Query Match 27.9%; Score 651.2; DB 12; Length 831;
Best Local Similarity 96.3%; Pred. No. 9.4e-131;
Matches 709; Conservative 0; Mismatches 23; Indels 4; Gaps 4;

QY 318 GTACCTGCTTCTGGAGCAGCTCTCGGGGGGTGAGCTATTCAGTACCTGTTAAGAGGG 377
Db 44 GTACCTGCTTCTGGAGCAGCTCTCGGGGGGTGAGCTATTCAGTACCTGTTAAGAGGG 103
QY 378 GAGAGTACGCCCAA-GGAGGCCGGAAGTCTTCCGCAATGTGTGCTGGCTGACT 436
Db 104 GAGAGTACGCCCAAAGGAGGCCGGAAGTCTTCCGCAAG-TGTGTGCTGGCTGACT 162
QY 437 TCTGCCACAGCTACTCCATCTGCCCACAGAGACTAAAGCCGAGAACCTGCTTTGGATG 496
Db 163 TCTGCCACAGCTACTCCATCTGCCCACAGAGACTAAAGCCGAGAACCTGCTTTGGATG 222
QY 497 AGAAAAACAATCCGATTCAGACTTGGAGCTTGGCATGCTCCCTGCAAGTGGGGACAGCC 556
Db 223 AGAAAAACAATCCGATTCAGACTTGGAGCTTGGCATGCTCCCTGCAAGTGGGGACAGCC 282
QY 557 TCTGTGAGACAGCTGCGGGGGGCCCCCAATATGCTGTCCAGAGGTGATTAAGGGGAAA 616
Db 283 TCTGTGAGACAGCTGCGGGGGGCCCCCAATATGCTGTCCAGAGGTGATTAAGGGGAAA 342
QY 617 AATATATGAGCCGCGGCGGAGACATGTGAGAGCTGTGAGTATCTCTTCCGCTGCTCG 676
Db 343 AATATATGAGCCGCGGCGGAGACATGTGAGAGCTGTGAGTATCTCTTCCGCTGCTCG 402
QY 677 TGGGGGCTTGGCTTTGATGAGACAACTTCCGCAAGCTGCTGAGAAAGTGAACGGG 736
Db 403 TGGGGGCTTGGCTTTGATGAGACAACTTCCGCAAGCTGCTGAGAAAGTGAACGGG 462
QY 737 GCGCTTCCACATGCCCCCACTTCAATCTCTCCAGATTCGATTCGAGAGCTCTGAGGGGAATGA 796
Db 463 GCGCTTCCACATGCCCCCACTTCAATCTCTCCAGATTCGAGAGCTCTGAGGGGAATGA 522
QY 797 TCGAAGTGAAGCCCGAAAAAGGCTCAGTGTGAGCAAAATTCAGAAACATCTTGTGTACC 856
Db 523 TCGAAGTGAAGCCCGAAAAAGGCTCAGTGTGAGCAAAATTCAGAAACATCTTGTGTACC 582
QY 857 TAGGCGGGAACACAGAGCCAGACCCGTGCTGAGACCAAGCCCTGGCCGCGGTAAGCA 916
Db 583 TAGGCGGGAACACAGAGCCAGACCCGTGCTGAGACCAAGCCCTGGCCGCGGTAAGCA 642
QY 917 TGGGG-AGCCCGGCATCCAAACGAGAGGTGAGACCCCGGAGCTGAGAGCATGAGCATCA 975
Db 643 TGGGGAGCCGCTCATCCAAACGAGAGGTGAG-CCCGAGCTCTAGAGAGCTGAGCATCA 701
QY 976 CTGGGCTGCTTTCAGAGGACCGCGAGAGGCTGATCGATCGAGCTGCGCAGTGAAGAGAGAAC 1035
Db 702 CTGGGCTGCTTTCAGAGGACCGCGAGAGGCTGATCGATCGAGCTGCGCAGTGAAGAGAGAAC 761
QY 1036 CAGAGAAAGATGATAT 1051
Db 762 ACAAAATGTTTATAT 777

RESULT 4
LOCUS BG819795 658 bp mRNA linear EST 22-MAY-2001
DEFINITION 602781839F1 NCL_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4932680
5', mRNA sequence.
ACCESSION BG819795
VERSION BG819795.1 GI:14167382

```

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE 1 (bases 1 to 658)
JOURNAL NIH-MGC http://mgi.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL0858 row: n column: 09
High quality sequence stop: 612.
Location/Qualifiers
1. 658
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4932680"
/clone_1lb="NCI-CGAP_Brn67"
/tissue_type="anaplastic oligodendroglioma with 1p/19q
loss"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCI-CGAP library."
BASE COUNT 145 a 194 c 223 g 96 t
ORIGIN
Query Match 25.4%; Score 602.2; DB 12; Length 658;
Best Local Similarity 98.4%; Pred. No. 3.7e-120;
Matches 619; Conservative 0; Mismatches 8; Indels 2; Gaps 1;
QY 859 GGGGGGAAACAGCAGCAGCCGCTGCTGAGCAGCCCTGCGCGCGGAGTGCATG 918
DB 1 GGGGGGAAACAGCAGCAGCCGCTGCTGAGCAGCCCTGCGCGCGGAGTGCATG 60
QY 919 CGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 978
DB 61 CGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY 979 GGGCTGCTGAGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1038
DB 121 GGGCTGCTGAGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
QY 1039 GAAAGATGAT 1098
DB 181 GAAAGATGAT 240
QY 1099 CAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1158
DB 241 CAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 1159 CGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1218
DB 301 CGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 1219 GATGCGGG 1278
DB 361 GATGCGGG 420
QY 1279 AGCCAGAGATCCCGTAGCGTCACTGAGAGCTCCACGGGCTGCTCTCCAGCCCTAAGC 1338
DB 421 AGCCAGAGATCCCGTAGCGTCACTGAGAGCTCCACGGGCTGCTCTCCAGCCCTAAGC 480
QY 1339 AGCCAGAGATCCCGTAGCGTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 1398

DB 481 AGCCAGAGATCCCGTAGCGTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 540
QY 1399 GGGGGGGGGTCCCGGAGCTTCCAAACAGCAGAGCTGCTGCTGCGGCGGAGTGCATG 1458
DB 541 GGGGGGGGGTCCCGGAGCTTCCAAACAGCAGAGCTGCTGCTGCGGCGGAGTGCATG 598
QY 1459 GGGGGGGGGGAGCAGCCCGGCGGCGGAGTGCATG 1487
DB 599 GGGGGGGGGGAGCAGCCCGGAGCAGAGAGAG 627
RESULT 5
LOCUS BM948524 698 bp mRNA linear EST 14-MAR-2002
DEFINITION UI-M-EG0p-bve-1-10-0-UI.r1 NIH_BMAP_EG0p Mus musculus CDNA clone
IMAGE:5691033 5', mRNA sequence.
ACCESSION BM948524
VERSION BM948524.1 GI:19432115
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 698)
JOURNAL NIH-MGC http://mgi.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pyX-5.
Location/Qualifiers
1. 698
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5691033"
/clone_1lb="NIH_BMAP_EG0p"
/tissue_type="whole brain"
/dev_stage="embryo 18.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/note="Organ: brain; Vector: pyX-Asc; Site: 1; EcoR I;
Site: 2; Not I; The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pyX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
, is CAGCCAGAC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institute of Mental Health
(NIMH), Hemlin Chin, Ph.D., program coordinator."
BASE COUNT 165 a 190 c 209 g 134 t
ORIGIN
Query Match 25.4%; Score 592.4; DB 14; Length 698;
Best Local Similarity 90.5%; Pred. No. 5e-118;
Matches 632; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 332 AGCAGCTCTGGGGGGGGTGAAGTATTCGACTACCTGTAAGAGGAGAGAGAGAGAGAG 391

```

||||| 1 AGCAGCTTTCGGTGTGAGCTGTTGCACTACCTGCTTAAAAAGGAGACTGACCCCA 60
||| 392 AGAGGCCCGGAAAGTTCTTCGCCGAGATGTGTGCGGTGAGCTTCCACAGCTACT 451
||| 61 AGAGGCCCGGAAAGTTCTTCGCCGAGATGTGTGCGGTGAGCTTCCACAGCTACT 120
||| 452 CCATCTGCGCAGAGAGCTTAAGCCCGGAGAACCTCTTTTGATGAGAAAAACATCC 511
||| 121 CCATCTGCGCAGAGAGCTTAAGCCCGGAGAACCTCTTTTGATGAGAAAAACATCC 180
||| 512 GCATTCGAGACTTGGCGATGCGTCCCTGCAAGTGGGGGACAGCTCTCGAGACAGCT 571
||| 181 GCATTCGAGACTTGGCGATGCGTCCCTGCAAGTGGGGGACAGCTCTCGAGACAGCT 240
||| 572 GCGGGTCCCGCATTTAGCGTGTCCAGAGTGTATTAAGGGGAAAAATATGATGGCCGC 631
||| 241 GCGGGTCCCGCATTTAGCGTGTCCAGAGTGTATTAAGGGGAAAAATATGATGGCCGC 300
||| 632 GGGGAGACATGTGAGCGTGTGAGTCACTCTTCCCTGCTGCGGGGCGCTTCCATGC 691
||| 301 GGGGAGACATGTGAGCGTGTGAGTCACTCTTCCCTGCTGCGGGGCGCTTCCATGC 360
||| 692 TTGATGACGACAACTCCGCGCAGCTGTGAGAGAGTGAACGGGCGCTTCCATGC 751
||| 361 TCGATGACGACAACTCCGCGCAGCTGTGAGAGAGTGAACGGGCGCTTCCATGC 420
||| 752 CCCACTTCATTCCTCCAGATTCGAGAGCTCTCTGAGGGGAAATGATGAAATGAGCCCG 811
||| 421 CTCACCTTCATTCCTCCAGATTCGAGAGCTCTCTGAGGGGAAATGATGAAATGAGCCCG 480
||| 812 AAAAAGGCTCAGTCTGAGAGAAATTCAGAAATCTTGTACTAGCCGGGAAACAG 871
||| 481 AGAAAGGCTCAGTCTGAGAGAAATTCAGAAATCTTGTACTAGCCGGGAAACAG 540
||| 872 AGCAGAGACCTGCTGAGAGCCAGCCCTGCGCGGGGTAGCAGTGCAGAGCTTCAT 931
||| 541 AACGAGACCTGCTGAGAGCCAGCCCTGCGCGGGGTAGCAGTGCAGAGCTTCAT 600
||| 932 CCAAGGAGAGCTGAGAGCCAGCCCTGCTAGAGAGAGTGCATGAGCTGAGGCTCTT 991
||| 601 CCAATGCGAGCTGAGAGCCAGCCCTGCTAGAGAGAGTGCATGAGCTGAGGCTCTT 660
||| 992 ACCGCGAGAGCTGAGAGCCAGCCCTGCTAGAGAGAGTGCATGAGCTGAGGCTCTT 1029
||| 661 ACCGCGAGAGCTGAGAGCCAGCCCTGCTAGAGAGAGTGCATGAGCTGAGGCTCTT 698

```

RESULT 6
 BM944329 742 bp mRNA linear EST 14-MAR-2002
 LOCUS BM944329
 DEFINITION U1-M-EHdp-bvq-d-01-0-01.r1 N1H_BMAP_EHdp Mus musculus cDNA clone
 IMAGE:5695440.5, mRNA sequence.
 ACCESSION BM944329
 VERSION BM944329.1 GI:19427914
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 742)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-t@mail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILND at:
 http://image.jnl.gov

This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: p1x-5.
 Location/Qualifiers
 1..742
 /organism="Mus musculus"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5695440"
 /clone_1db="N1H_BMAP_EHdp"
 /tissue_type="whole brain"
 /dev_stage="embryo 18.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /note="Organ: brain; Vector: p1x-asc; Site: 1: EcoR I;
 Site: 2: Not I; The library was constructed according to
 Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into p1x-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is CAGCCAGCAGC. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP). Gene discovery in the Developing Mouse Nervous
 System, supported by National Institute of Mental Health
 (NIMH), Hemin Chin, Ph.D., Program coordinator."

BASE COUNT 179 a 197 c 219 g 145 t 2 others
 ORIGIN

Query Match 24.6%; Score 574.8; DB 14; Length 742;
 Best Local Similarity 89.4%; Pred. No. 3.4e-114;
 Matches 663; Conservative 0; Mismatches 74; Indels 5; Gaps 4;

```

442 CACGCTACTTCATCTGCGCAGAGAGCTTAAGCCCGGAGAACCTCTTTGATGAGAGAA 501
||| 2 CATAGCTACTTCATCTGCGCAGAGAGCTTAAGCCCGGAGAACCTCTTTGATGAGAGAA 61
||| 502 AACAACTCCGATTCGAGAGCTTGGCATGCGTCCCTGCAAGTGGGGGACAGCTCTCTG 561
||| 62 AACAACTCCGATTCGAGAGCTTGGCATGCGTCCCTGCAAGTGGGGGACAGCTCTCTG 121
||| 562 GAGACAGCTGCGGGTCCCGCATTTAGCGTGTCCAGAGTGTATTAAGGGGAAAAATAT 621
||| 122 GAGACAGCTGCGGGTCCCGCATTTAGCGTGTCCAGAGTGTATTAAGGGGAAAAATAT 181
||| 622 GATGGCGCGCGGAGAGATGTGAGAGCTGTGAGAGCATCTCTTCCCTGCTGCTGGGG 681
||| 182 GATGGCGCGCGGAGAGATGTGAGAGCTGTGAGAGCATCTCTTCCCTGCTGCTGGGG 241
||| 682 GCTCTGCGCTTTGATGAGAGAGCTTCCCGCAGCTGCTGAGAGAGTGAAGCGGGCGTC 741
||| 242 GCACTGCGCTTTCATGAGAGAGAGCTTCCCGCAGCTGCTGAGAGAGTGAAGCGGGCGTC 301
||| 742 TTCCACATGCGCCCATTCATCTTCAGAGATTCAGAGAGCTTCTGAGGGGAAATGATG 801
||| 302 TTCCACATGCGCCCATTCATCTTCAGAGATTCAGAGAGCTTCTGAGGGGAAATGATG 361
||| 802 GTGAGAGCGCGGAAAAAGGCTCAGTGTGAGAGCAATTCAGAAACATCTTGTGA--CTAG 859
||| 362 GTGAGAGCGCGGAAAAAGGCTCAGTGTGAGAGCAATTCAGAAACATCTTGTGA--CTAG 421
||| 422 GCGGGAACACAGAGACAGCCGCTGAGAGAGCCCTGCGCGCGGAGTGGAGCATATGC 919
||| 919 GCGGGAACACAGAGACAGCCGCTGAGAGAGCCCTGCGCGCGGAGTGGAGCATATGC 481
||| 482 GTACCTGCTTCCTCAAGGAGAGAGCTGAGAGAGCTTCTTGAAGAGCATATGAGTCTGG 541
||| 541 GTACCTGCTTCCTCAAGGAGAGAGCTGAGAGAGCTTCTTGAAGAGCATATGAGTCTGG 980
||| 980 GTACCTGCTTCCTCAAGGAGAGAGCTGAGAGAGCTTCTTGAAGAGCATATGAGTCTGG 1039

```

Db 542 GCTGCTTCAGAGACCGGAGCGCTGCACAGAACTGCGAAGGAGAGAAAACCAAG 601
 QY 1040 AAAAGATGATATATATATATGTC-TTTTGATCGAAGAGCGGTATCCAGCTGTAGAGAC 1098
 Db 602 AAAAGATGATATATATATATGTC-TTTTGATCGAAGAGCGGTATCCAGCTGTAGAGAC 661
 QY 1099 CAGAGCTGCTCCCGGAAATGATGTTGACC-CCCCGGAGAGGTGTGATTTCCCAT 1157
 Db 662 CAGGA-CTGCTCTCTCGGAAATGATGTTGACCCTCGTAGCGGTGTGATTTCCCAT 720
 QY 1158 GCTGAGCCGTCAGGGAAGCG 1179
 Db 721 GCTGACCAACAGCGGAACGG 742

RESULT 7
 B0178547 671 bp mRNA linear EST 30-APR-2002
 LOCUS B0178547
 DEFINITION UT-M-EV0-bvs-b-03-0-UI.r1 NIH-BMAP-EV0 Mus musculus cDNA clone
 IMAGE:5701154 5', mRNA sequence.
 ACCESSION B0178547
 VERSION B0178547.1 GI:20354039
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 671)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
 http://image.lnll.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES
 source
 Seq primer: pyx-5.
 Location/Qualifiers
 1..671
 /organism="Mus musculus"
 /strain="C57Bl/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5701154"
 /clone_1lb="NIH-BMAP-EV0"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /note="Organ: brain; Vector: pyx-Asc; site:1: EcoR I;
 site:2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pyx-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is GTGCGTGA. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP). 'Gene discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemm Chlin, Ph.D., program coordinator."

BASE COUNT 160 a 188 c 200 g 122 t 1 others
 ORIGIN
 Query Match 23.3%; Score 543; DB 14; Length 671;
 Best Local Similarity 90.0%; Pred. No. 2.6e-107;

Matches 604; Conservative 0; Mismatches 65; Indels 2; Gaps 2;
 QY 772 TGCCAGAGCTCC-TGAGGGAGATGATCAAGTGGAGCCGGAAGAAAGGCTCACTGGA 830
 Db 1 TGCCAGAGCTCCNTGAGAGGATGATGAAAGTGGAGCCGGAAGAAAGGCTCACTGGA 60
 QY 831 GCAAATTCAGAAACATCTTGTGATGAGCGGGGAAACAGAGCAGACCCGCTGGA 890
 Db 61 GCAAAATTCAGAAACATCTTGTGATGAGCGGGGAAACAGAGCAGACCCGCTGGA 120
 QY 891 GCCAGCCCTGCGCGCGGTAGCCATGCGAGCCTCCATCCAGAGAGCTGAGACC 950
 Db 121 GCCAGCCCGAGCGAGATGACCATGCGTACCTGCTTCAGATGGCAGCTGAGACC 180
 QY 951 CGACGCTCTAGAGAGCATGACATGAGCGCTGCTTCCAGAGCCGAGAGGCTGATCG 1010
 Db 181 TGACGTTCTGGAAGATGAGCTGCTGCTGCTGCTTCCAGAGCCGAGCGCTGACAG 240
 QY 1011 CGAGCTGCGCAGTGGAGAGAGAAACCAAGAAAGATGATATATGCTTTGATCG 1070
 Db 241 AGAACTGCGAAGCGAGAGAGAAACCAAGAAAGATGATATATGCTTTGATCG 300
 QY 1071 GAAAGAGCGGTATCCAGCTGTGAGAGAC-AGAACCTGCTCCCGGAATGATGAC 1129
 Db 301 GAAAGAGCGGTATCTAGCTGTGAAAGCAAGAGCTGCTCTCGGAATGATGATGAC 360
 QY 1130 CCCCCGGAAGCGTGTGATCTCCATGCTGAGCCGTCACGGAGCGGAGCAGAGC 1189
 Db 361 CACCTGTGAAGCTGTGATGATGCTCCCATGCTGAGCAGACAGGAAACGAGCGCAAGC 420
 QY 1190 GGAAGTCATGGAAGTCTGAGCATCACGATGCCGGGGGTGTGCTCCCTGATCCA 1249
 Db 421 GGAAGTCATGGAAGTCTGAGCATCACGATGCTGGAGTGTGCTCCCATGATGCTA 480
 QY 1250 CCGAGAGGCGCTGAGATGCGCCAGACAGCAAGATCCGATAGCTAGTGAACCT 1309
 Db 481 CCGAGAGGCGCTGAGATGCGCCAGACAGCAAGATCCGATAGCTAGTGAACCT 540
 QY 1310 CCACGGCTGTGCTCTCAGGCTCTAAGCAGCCCAAGAGTCCGCTTCTTTCAC 1369
 Db 541 CCACGGCTGTGCTCTCAGGCTCTAAGCAGCCCAAGAGTCCGCTTCTTTCAC 600
 QY 1370 CGAGAGCGGGGCTGAGATGAGCTGAGCGGGGGCTCCCGACTCCAAAACGAGA 1429
 Db 601 CGAGAGCGGGGCTGAGATGAGCTGAGCGGGGGCTCCCGACTCCAAAACGAGA 660
 QY 1430 CGCTGCTCT 1440
 Db 661 CGCTGCTCT 671

RESULT 8
 B0573641 658 bp mRNA linear EST 19-JUN-2002
 LOCUS B0573641
 DEFINITION UT-M-EV0-bvs-b-15-0-UI.r1 NIH-BMAP-EV0 Mus musculus cDNA clone
 IMAGE:5718158 5', mRNA sequence.
 ACCESSION B0573641
 VERSION B0573641.1 GI:21476958
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 658)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa


```

|||||
Db 305 TACTGTGCTAGAACACACCTGTGACGCTTTCGACTTGTGTAAGAGGCG 364
OY 379 AGACTGAGCCCAAGAGAGGCGGAAAGTTCCTCCGAGATTGTCTGCGCTGACCTTC 438
Db 365 AGGCTGACCCCTAAGAGAGGCTCGGAAGTTCCTCCGAGATCATCTGCGCTGACCTTC 424
OY 439 TGGCAGAGCAATCCATCTGTCACAGAGACCTTAAAGCCCGAGAACCTGCTTTGATGAG 498
Db 425 TGGCAGAGCAATCCATCTGTCACAGAGATCTTAAAGCCCGAGAACCTGCTTTGATGAG 484
OY 499 AAAAACAACATCCGATCTGACACTTCCGATGCGCTCCCTGTCAGAGTGGGAGACAGCTTC 558
Db 485 AAGAACAACATCCGATCTGACACTTCCGATGCGCTCCCTGTCAGAGTGGGAGACAGCTTC 544
OY 559 CTGAGACCAAGCTGCGGCTCCCGCATTTATGCTGTCAGAGAGTATTAAGGGGAGAAAA 618
Db 545 TTGGAGACCAAGCTGCGGCTCCCGCATTTATGCTGTCAGAGAGTATTAAGGGGAGAAAA 604
OY 619 TATGATGGCGCGCGGAGACATGTCGAGCTGTGAGTCAATCTTGTGCGCTGCTCTG 678
Db 605 TATGACGCGCGG-GAAGGAGAGCTGTGAGCTGCGGCTGTCATCTGCTGCTGCTG 663
OY 679 GGGGCTCTGCGCTTGTGATGACAGACACTCCGCGAGCTGTGAGAGAGTAAAGGAGGCG 738
Db 664 GGGGCTCTGCGCTTGTGATGACAGACACTCCGCGAGCTGTGAGAGAGTAAAGGAGGCG 722
OY 739 GTCTTCCACATGCCCCACTTATCTTCCAGATTGCCAGAGCTCTGAGGGGAAATGATC 798
Db 723 GTCTTCCACATGCCCCACTTATCTTCCGCGCGAGCTGTGAGAGTCTGCTGAGGAGATGATC 782
OY 799 GAAGTGAGGCGCGGAGAAAAAGGCTGTGAGCAATTCAGAAACATCTCTGG 852
Db 783 GAGGTGAGCGCGGAGCGCGCGCTCAAGCTAGAGCAATTCAGAAACATATATG 836

RESULT 10
BF529630 936 bp mRNA linear EST 11-DEC-2000
LOCUS 602043855F1 NCL_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4181581
DEFINITION 5', mRNA sequence.
ACCESSION BF529630
VERSION BF529630.1 GI:11616993
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 936)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.lnl.gov
Plate: LAM9494 row: n column: 14
High quality sequence start: 8
High quality sequence stop: 568.
Location/Qualifiers
1. 936
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4181581"
/clone_id="NCL_CGAP_Brn67"
/tissue_type="anaplastic oligodendroglioma with 1p/19q
loss"
/lab_host="DH10B (T1 phage-resistant)"

```

```

/Note="Organ: brain; Vector: pCMV-Sport6; Site 1: NCI;
Site 2: Sall; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCL_CGAP Library."
BASE COUNT 286 a 238 c 318 g 94 t
ORIGIN
Query Match 22.4% Score 522.4; DB 12; Length 936;
Best Local Similarity 97.6%; Pred. No. 8.2e-103;
Matches 562; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

OY 901 GGGCGCGGAGTAGCAGCAGGAGCGGAGCTGCAATCCAGAGAGAGCTGGAGCCGAGCTGCTA 960
Db 5 GGGCGCGGAGTAGCAGCAGGAGCGGAGCTGCAATCCAGAGAGAGCTGGAGCCGAGCTGCTA 64
OY 961 GAGAGCATGGCATCACTGGGCTGCTTCAAGGAGCCGAGAGGCTGCATGCGAGCTGCGC 1020
Db 65 GAGAGCATGGCATCACTGGGCTGCTTCAAGGAGCCGAGAGGCTGCATGCGAGCTGCGC 124
OY 1021 AGTGAGAGAGAGAAACCAAGAAAGATATATATATCTGCTTTTGGATCGAAGAGCGG 1080
Db 125 AGTGAGAGAGAGAAACCAAGAAAGATATATATATCTGCTTTTGGATCGAAGAGCGG 183
OY 1081 TATCCAGCTGTGTGAGAGACAGAGACCTGCTCCCGGAGATGATGTGACCCCGCGAGAG 1140
Db 184 TATCCAGCTGTGTGAGAGACAGAGACCTGCTCCCGGAGATGATGTGACCCCGCGAGAG 243
OY 1141 CGTGTGATTTCTCCATGCTGAGCGCTACAGGAGGCGGAGCGGAGCGGAGTCTATG 1200
Db 244 CGTGTGATTTCTCCATGCTGAGCGCTACAGGAGGCGGAGCGGAGCGGAGTCTATG 302
OY 1201 GAAGTCTGAGCATCACTGGGCTGCTTCAAGGAGCCGAGAGGCTGCATGCGAGCTGCGC 1260
Db 303 GAAGTCTGAGCATCACTGGGCTGCTTCAAGGAGCCGAGAGGCTGCATGCGAGCTGCGC 362
OY 1261 TTGAGATGGCGCGAGCAGACAGAGATCCGATGAGCTGAGAGCTTCCAGGAGCTG 1320
Db 363 TTGAGATGGCGCGAGCAGACAGAGATCCGATGAGCTGAGAGCTTCCAGGAGCTG 421
OY 1321 TCTTCCAGCCCTGTAGCAGCCCAAGAGTCCGCTTTTCTTTTCAACCGAGCGGGG 1380
Db 422 TCTTCCAGCCCTGTAGCAGCCCAAGAGTCCGCTTTTCTTTTCAACCGAGCGGGG 481
OY 1381 GCTGAGATGAGGCTGAGGCGGCGGCTCCCGGAGCTTCCAGAAAGCAGAGCTGCTTCT 1440
Db 482 GCTGAGATGAGGCTGAGGCGGCGGCTCCCGGAGCTTCCAGAAAGCAGAGCTGCTTCT 541
OY 1441 CGGGGCGCGCAGGGGTGGGCGCGCGGAGCAGCGCC 1476
Db 542 TCGGGGCGCGCAGGGGTGGGCGCGCGGAGCAGCGC 577

RESULT 11
B0086295 510 bp mRNA linear EST 29-APR-2002
LOCUS 1120b01.y1 Melton Normalized Human Islet 4 M-HIS 1 Homo sapiens
DEFINITION cDNA clone IMAGE:6135408 5' similar to TR:060843 060843 PUTATIVE
SERINE/THROMBIN PROTEIN KINASE ;, mRNA sequence.
ACCESSION B0086295
VERSION B0086295.1 GI:20045499
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 510)
REFERENCE Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Secaese, M., Brestelli, J., Gridwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Rittler, E., Ronko, I., Bennett, D., Cardenas
M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T.,
Jackson, Y., and Bowers, Y.
Endocrine Pancreas Consortium
TITLE

```


Qy	558	CCTGGAGACAGACGTGCGGGTCCCCCAATTATGCGTGCACAGCGATTAAGAGGGGAAAA	617
Db	329	CCTGGAGACAGACGTGCGGGTCCCCCAATTATGCGTGTCCAGAGCGATTAAGGGGGAAAA	388
Qy	618	ATATGATATGCGCGCGGGCGACACATGTGGAGCTGTGAGATCACTCTCTTCCGCCCTGCTGT	677
Db	389	ATATGATATGCGCGCGGGCGACACATGTGGAGCTGTGAGATCACTCTCTTCCGCCCTGCTGT	448
Qy	678	GGGGCTGTGCCCTTTGTATGACGACAACTCCGCGACGTGCTGGAGAGTGAACGGAACGGGG	737
Db	449	GGGGCTGTGCCCTTTGTATGACGACAACTCCGCGACGTGCTGGAGAGTGAACGGAACGGGG	508
Qy	738	CGTCTTCACATGCCCCACCTTCATTTCTCCAGATTTGCCAGAGCTCTGTGAGGGGAAAT--G	795
Db	509	CGTCTTCACATGCCCCACCTTCATTTCTCCAGATTTGCCAGAGCTCTCTGAGGGGAAATTGA	568
Qy	796	ATCGAATGAGAGCCCG--AAAAAAGGCTAGTCTGGAGCAATTCAGAA	842
Db	569	TCGAAATGAGAGCCCGAAAAAAGGCTAGTCTGAGCCCAATTCGAA	616

RESULT 13					
B0444032					
LOCUS					
DEFINITION					
IMAGE:5708798 5', mRNA sequence.					
	600 bp	mRNA	linear	EST 29-MAY-2002	
B0444032					
UT-M-EMO-bx1-p-15-0-UI-r1 NIH-BMAP-EMO					
			Mus musculus	CDNA clone	

ORGANISM	Mus musculus
SOURCE	house mouse.
KEYWORDS	EST.
VERSION	BQ444032.1
ACCESSION	BQ444032
GI	21247144

REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia, Sclurognathi, Muridae, Mus
AUTHORS	1 (bases 1 to 600)
TITLE	NIH-MGC http://mgc.nhl.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D.

Email: cgaabos-remail.ain.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pyx-5.

```

FEATURES
SOURCE
Location/Qualifiers
1. .600
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone_image="5708798"
/clone_id="NIH_BMAP_EW0"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I;
Site_2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
, is GTGGCGGAA. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP). 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health

```

BASE COUNT	(NIMH) /	Hemin Chln,	Ph.D.,	Program	coordinator.."
145 a	156 c	173 g	125 t	1	others
ORIGIN					

Query Match	21.08;	Score 490.2;	DB 14;	Length 600;
Best Local Similarity	91.48;	Pred. No. 7e-96;		
Matches 541;	Conservative 0;	Mismatches 49;	Indels 2;	Gaps 2;

274 CATGTCCTCAAGCTCCACGACGCTACGAGMACAAGAAATTTGTACTCGTTCGGAG 333

Db 11 CACGTGCTCAAGCTCCACGACGCTACGAGAACAGAAATATTATACTTGGTCTTGAG 70

334 CACGTCCTGGGGGTGAGCTATTGCACTACCTGCTAAAGAAGGGGAGAGACTGACGCCCAAG 393

Db 71 CACGTTTCGTGGTGAGCTGTTGACTACCTGGTAAAAAAGGAGACTGACACCCAAG 130

394 GAGCCCCAAGTCTTCCGCAAGATTGTGTCGGCGGACTTCTGCCAAGCTACTCC 453

[illegible]

DJ
L01 9800CCCCCHT01LC1ICCCBCCA01CGI9IGCBPC1000FC1C100CAIN00CIACILCL LSV

454 ATCTGCCACAGACCTAAAGCCCGAGACCTGCTTTTGGATGAGAAAACAACATCCGC 513

Db 191 ATCTGTCACAGAGACTTGAAAGCCAGAGAACCCTGCTGTTGGATGAGAAAAACAACATCCGC 250

514 ATTGCAGACTTCGGCATGGCGTCCCTGCAGGTGGGGGACAGCCCTCTGGAGACCAGCTGC 573

Db 251 ATGCAGACTTTGGTATGGCGTCTCCTGCAAGTGGGGACAGAGCTCTCTGGAGACCAAGCTGC 310

574 GGGTCCCCCATTTATGCGTGTCCAGAGGTGATTTAAGGGGAAAAATATGATGGCCGCGG 633

Db 311 GGGTCCCCCAATTACGCANGTCCAGAGGTGATCAAGGGGGAAAAGTATGATGGCCCGCGG 370

[illegible][illegible]

3 / 1 GCAGACATGTCGAGCCGTGTGAGTTCATTCCTATTGCCCTGCTTGTGTGGGGCAGCTGCCCTTC 430

694 GATGACGACAACCTCCGCCAGCTGCTGGAGAAAGGTGAACCGGGCGCTCTCCACATGCCC 753

Db 431 GATGACGAC-ACCTGGCCAGCTACTGGAGAGGTGAACGCTGGGGTCTTCCACATGCCCT 489

QY 754 CACTTCATTCCTCCAGATTGCCAGACCTCCTGAGGGGATGATCGAAGTGGAGCCCGA 813

Db 490 CACTTCATCCCTCCAGACTGCCAGGCTCTGAGAGGATGATTGAAGTGGAGCCCGAG 549

QY 814 AAAAGGCTCAGTCTGGAGCAATTCAGAACATCCTTGTTACCTAGGCGGGA 865

550 A A A A G G C T C A G T T C T G G A G C - A A T T C A G A A A C A T C C C T G C A T T C T G G C G C G C G A 600

[illegible]

RESULT 14

BF796285	753 bp	mRNA	linear	EST 12-JAN-2001
LOCUS	BF796285			

DEFINITION 602258590F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4341838 5', mRNA sequence

INSTRUMENTS REQUIRED:
ACCESSION BF796285
 00700005 1 01 10101000

VERSION	BF/96285.1	G1:12101339
KEYWORDS	EST.	

SOURCE	human.
ORGANISM	Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo

REFERENCE
1 (bases 1 to 753)

AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis Staudt, M.D., PhD

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.S.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/JINP at:

http://image.llnl.gov
Plate: LHM956 row: c column: 23
High quality sequence stop: 691.

FEATURES

source

1. 753
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4341838"
/issue="1b"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph. Vector: pCMV-Sport6; Site: 1: NotI; Site 2: SalI. Cloned unidirectionally. oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."

BASE COUNT 156 a 212 c 262 g 123 t
ORIGIN

Query Match 20.9%; Score 487.8; DB 12; Length 753;
Best Local Similarity 96.8%; Pred. No. 2.5e-95;
Matches 572; Conservative 0; Mismatches 12; Indels 7; Gaps 7;

QY 858 AGGGGGAACAGAGGACGAGCCGCTGGAGCCAGCCCTGGCCGGGATGACCAT 917
DB 167 AGGGGGAACAGAGGACGAGCCGCTGGAGCCAGCCCTGGCCGGGATGACCAT 226
QY 918 GCGGAGCCTGCATCAACAGAGAGCTGAGCCCGAGCTCTAGAGAGCATGGCATCT 977
DB 227 GCGGAGCCTGCATCAACAGAGAGCTGAGCCCGAGCTCTAGAGAGCATGGCATCT 286
QY 978 GGGCTGCTTCAGGAGCGCGAGAGCTGCATCGGAGCTGCGCAGTGAAGAGAGCA 1037
DB 287 GGGCTGCTTCAGGAGCGCGAGAGCTGCATCGGAGCTGCGCAGTGAAGAGAGCA 346
QY 1038 AGAAAGATGATATATATATCTGCTTTGGATCGGAAGAGCGGTATCCAGCTGTAGGA 1097
DB 347 AGAAAGATGATATATATCTGCTTTGGATCGGAAGAGCGGTATCCAGCTGTAGGA 406
QY 1098 CAGAGACCTGCTCCCGGAGATGATGTGACCCCGCGAGAGCTGTGATTTCCCAT 1157
DB 407 CAGAGACCTGCTCCCGGAGATGATGTGACCCCGCGAGAGCTGTGATTTCCCAT 466
QY 1158 GCTGAGCCGTACGAGGAGCGCGAGAGCTGCATCGAAGTCTTGAAGCATCAC 1217
DB 467 GCTGAGCCGTACGAGGAGCGCGAGAGCTGCATCGAAGTCTTGAAGCATCAC 525
QY 1218 CGATGCGGGGGGTGGTGGCTCCCTGTACCCAGCGGAGCGCTTGGAGATGGCCAGCA 1277
DB 526 CGATGCGGGGGGTGGTGGCTCCCTGTACCCAGCGGAGCGCTTGGAGATGGCCAGCA 585
QY 1278 CAGCAGAGATCCCGTACGTCAGT-GGAGGCTCCAGCGGCTGTCTCCAGCCCTCTAA 1336
DB 586 CAGCAGAGATCCCGTACGTCAGT-GGAGGCTCCAGCGGCTGTCTCCAGCCCTCTAA 644
QY 1337 GCAGCCCAAGAGATCGCTGCTTTTCTTTTCAACGAGCGCGGGGCTGGAGATGAGGCTC 1396
DB 645 GCAG-CCAAAGAGTCCGCTCTTTTCTTTTCAACGAGCGCG-GGGGCTGGAAGATGAGGCTC 702
QY 1397 GAGCGGGGGGCTCCCGCATCTCCAAAGAGAGCTGCTTCTGGGGGCC 1447
DB 703 GA-CCGGGGGCTCCCGCA-TTCCAAAGAGAGAGCTGCTTCTGGGGGCC 751

RESULT 15 614 bp mRNA linear EST 14-MAR-2002
BM944236
LOCUS BM944236
DEFINITION UT-M-EHOP-bvq-a-22-0-UT.r1 NIH_BMAP_EHOP Mus musculus cDNA clone
IMAGE:5695389 5, mRNA sequence.
ACCESSION BM944236
VERSION BM944236.1 GI:19427821
KEYWORDS EST.
SOURCE house mouse.

ORGANISM

Mus musculus

REFERENCE

Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS

1 (bases 1 to 614)

TITLE

NIH-MGC http://imgc.nci.nih.gov/

JOURNAL

National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pyx-5.

FEATURES

source

Location/Qualifiers

1. 614
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5695389"
/issue="1b"
/dev_stage="embryo 18.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/note="Organ: brain; Vector: pYX-asc; Site: 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-asc vector. The library tag sequence located between the Not I site and the polyA tail, is CAGCAGCAGC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). Gene discovery in the developing Mouse Nervous System, supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

BASE COUNT

153 a 165 c 183 g 113 t

Query Match

20.4%; Score 476.2; DB 14; Length 614;

Best Local Similarity

89.1%; Pred. No. 7.6e-93;

Matches 548; Conservative

0; Mismatches 63; Indels 4; Gaps 3;

QY 440 GCCACAGTACTCATCTGCACAGAGACCTAAACCCGAGAACCTGCTTTGATGAGA 499
DB 1 GCCATAGCTATCTCATCTGTCACAGAGATGTAAG-CAGAGAACCTGCTTTGATGAGA 59
QY 500 AAAACAACATCCGATTCGACAGTTCGCGATGGGCTCCCTGAGGTGGGGACAGCTCC 559
DB 60 AAAACAACATCCGATTCGACAGTTCGCGATGGGCTCCCTGAGGTGGGGACAGCTCC 119
QY 560 TGGAGACAGCTGGGGGCCCCCATTTATGGTGCAGAGGTGATTAAGGGGAAAAAT 619
DB 120 TGGAGACAGCTGGGGGCCCCCATTTATGGTGCAGAGGTGATTAAGGGGAAAAAT 179
QY 620 ATGATGCGCGCGGAGACATGTGAGCTGTGAGTATCTTTCGCCCTGCTGCTGG 679
DB 180 ATGATGCGCGCGGAGACATGTGAGCTGTGAGTATCTTTCGCCCTGCTGCTGG 239
QY 680 GGGCTGCGCTTTGATGAGACAACTTCGCCAGCTGCTGGAGAGTGAACCGGGGCG 739
DB 240 GGGCAGTCCCTTGATGAGACAACTTCGCCAGCTGCTGGAGAGTGAACCGGGGCG 299
QY 740 TCTTCACATCCCGCATCTCATCTCCAGATGTCAGAGCTCCTGAGAGGAATGATCG 799

Db	300	TCCTCCACATGCGCTCACTTCATCCCTCCACAGACTGCCAGAGCGCTCTAGAGGGATGATG	359
QY	800	AAGTGAGGCCCGCAAAAAGAGCTAGTCTGGAGCAAAATTCAAGAAACA--TCCTTGGTACT	857
Db	360	AAGTGAGGCCCGAGAAAAGGCTAGTCTGGAGCAAAATTCAAGAAACAATCCCTGGTATCTGT	419
QY	858	AGGCGGGAAACAAGAGCC--AGACCCCTGCTCTGGAGGCCGCCCTCGGCCGCGGTAGCA	916
Db	420	GGCGGGGAAACAAGAACCAAGACCTTGGCTGGAGCCAGGCCCGGCGCGAGATGAGCA	479
QY	917	TGCGGAGCGCTGCATCCAAAGAGAGACTGGACCCCGACGTCCTTAGAAGCATGGCATAC	976
Db	480	TGCGTAGCGCTGCCCTTCACATGGGAGACTGGACCTGACGTTCTGGAAAAGCATGGCGTCTC	539
QY	977	TGGGCTGCTTCAGGAGCACCGAGAGAGGCTGCATCGCGAGCTGGCGCATGAGAGAGAGAACC	1038
Db	540	TGGGCTGCTTCAGAGAGCACCGAGAGGCGCTGCACAGAGAAATCGCGAAGGAGAGAGAAACC	599
QY	1037	AAGAAAGATGATAT 1051	
Db	600	AAGAAAGATGATCT 614	

Search completed: April 22, 2003, 00:23:49
Job time : 2137.25 secs

GenCore version 5.1.4-P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 19:23:30 ; Search time 320.629 Seconds

(without alignments)
16393.314 Million cell updates/sec

Title: US-10-003-690-3

Perfect score: 2334
Sequence: 1 atgtctccggggcgaaga.....ccaacggagccctctgcc 2334

Scoring table: IDENTITY_MDC
Gap 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID2/gcgdata/geneseq/emb1/NA1980.DAT:*
2: /SID2/gcgdata/geneseq/emb1/NA1981.DAT:*
3: /SID2/gcgdata/geneseq/emb1/NA1982.DAT:*
4: /SID2/gcgdata/geneseq/emb1/NA1983.DAT:*
5: /SID2/gcgdata/geneseq/emb1/NA1984.DAT:*
6: /SID2/gcgdata/geneseq/emb1/NA1985.DAT:*
7: /SID2/gcgdata/geneseq/emb1/NA1986.DAT:*
8: /SID2/gcgdata/geneseq/emb1/NA1987.DAT:*
9: /SID2/gcgdata/geneseq/emb1/NA1988.DAT:*
10: /SID2/gcgdata/geneseq/emb1/NA1989.DAT:*
11: /SID2/gcgdata/geneseq/emb1/NA1990.DAT:*
12: /SID2/gcgdata/geneseq/emb1/NA1991.DAT:*
13: /SID2/gcgdata/geneseq/emb1/NA1992.DAT:*
14: /SID2/gcgdata/geneseq/emb1/NA1993.DAT:*
15: /SID2/gcgdata/geneseq/emb1/NA1994.DAT:*
16: /SID2/gcgdata/geneseq/emb1/NA1995.DAT:*
17: /SID2/gcgdata/geneseq/emb1/NA1996.DAT:*
18: /SID2/gcgdata/geneseq/emb1/NA1997.DAT:*
19: /SID2/gcgdata/geneseq/emb1/NA1998.DAT:*
20: /SID2/gcgdata/geneseq/emb1/NA1999.DAT:*
21: /SID2/gcgdata/geneseq/emb1/NA2000.DAT:*
22: /SID2/gcgdata/geneseq/emb1/NA2001A.DAT:*
23: /SID2/gcgdata/geneseq/emb1/NA2001B.DAT:*
24: /SID2/gcgdata/geneseq/emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2256	96.7	2385	22	AAS06717
2	2256	96.7	2897	24	AA026464
3	961	41.2	2025	24	ABA02995
4	961	41.2	2217	24	ABA02994
5	955.8	41.0	2647	24	AA034315
6	658.8	28.2	906	22	ABA08296
7	614	26.3	614	22	ABA08924
8	512	21.9	512	24	ABK70216
9	470.8	20.2	2720	23	ABL04689

10	264.2	11.3	2954	22	AAK94923	Human full-length
11	264.2	11.3	3269	22	AA018831	Human kinase (PKIN
12	264.2	11.3	3312	24	ABA05739	Human neuronal ser
13	264.2	11.3	3392	24	ABA05740	Human neuronal ser
14	261.6	11.2	1594	22	AA044655	Novel protein kina
15	259.4	11.1	3170	24	ABA05737	Murine neuronal se
16	259.4	11.1	3250	24	ABA05738	Murine neuronal se
17	257.8	11.0	316	22	ABA13768	Human neuronal se
18	257.8	11.0	316	22	AA027156	Human neuronal se
19	257.8	11.0	316	22	ABA03545	Human neuronal se
20	240.6	10.3	2112	24	AA030397	DNA encoding nove
21	240.6	10.3	2222	24	AA030398	DNA encoding nove
22	240.6	10.3	2715	22	AAK51482	Human PAR-1B beta
23	238.8	10.2	2663	23	ABA05735	Human polynucleoti
24	237.8	10.2	2979	22	AAK52456	Human polynucleoti
25	229	9.8	4354	23	ABL07437	Drosophila melanog
26	221.4	9.5	2352	22	AA068823	Human protein kina
27	221.4	9.5	2533	21	AAK82951	Human keratinocyte
28	221.4	9.5	2632	21	AAK82952	Human keratinocyte
29	221.4	9.5	2968	22	AA068822	Human keratinocyte
30	221.4	9.5	4699	22	AAH76213	Human kinase PKIN-
31	212.8	9.1	1536	14	AA047789	SHP gene. Yeast
32	212.4	9.1	2222	24	AA033068	Human PAR-1B beta
33	211.2	9.0	2118	23	ABL11605	Drosophila melanog
34	211.2	9.0	4118	23	ABL11604	Drosophila melanog
35	207	8.9	2766	23	ABL26615	Drosophila melanog
36	207	8.9	3957	21	AA074495	Human ORFX ORF3050
37	207	8.9	4321	22	AA044652	Novel protein kina
38	207	8.9	4381	24	AA034304	Human PKIN-7 cDNA.
39	206.6	8.9	1647	18	AA059925	Mammalian AMPK alp
40	204	8.7	1349	22	AA199637	Human expressed po
41	204	8.7	1349	22	ABK43708	DNA encoding novel
42	202.4	8.7	2361	22	AAS06718	Polynucleotide seq
43	201	8.6	859	22	AAK91635	Human cDNA 5'-end
44	201	8.6	859	22	AAK93561	Human cDNA clone r
45	201	8.6	2024	22	AAK94157	Human full-length

ALIGNMENTS

RESULT 1
ID AAS06717 standard; cDNA; 2385 BP.
AC AAS06717;
XX
XX
DT 12-SEP-2001 (first entry)
XX
XX Polynucleotide sequence encoding human protein kinase #17.
DE
XX
XX Human; protein kinase; PK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder; gene therapy; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200138503-A2.
XX
XX
XX PD 31-MAY-2001.
XX
XX
XX PE 22-NOV-2000; 2000MO-US32085.
XX
XX
XX PR 24-NOV-1999; 99US-0167482.
XX
XX (SUGF-) SUGEN INC.
XX
XX Plozman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
XX PI Flanagan P, Clary D;
XX
XX WPI: 2001-343950/36.
XX
XX P-PSDB; AA003517.

XX Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections -
PS Example 1: Figure 1: 433pp; English.

CC AAS06701-AAS06757 encode for novel human protein kinases #1-57. The
CC novel protein kinases have been identified as members of the tyrosine
CC or serine/threonine kinase (PRK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of hematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be
CC used for gene therapy and as DNA probes in diagnostic assays.
CC The protein kinase polypeptides may be used as antigens in the production
CC of antibodies against the protein kinases and in assays to identify
CC modulators of protein kinase expression and activity.

XX Sequence 2385 BP; 494 A; 768 C; 731 G; 392 T; 0 other;

Query Match 96.7%; Score 2256; DB 22; Length 2385;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 CAGGCCCAATATGTGGCCCTATCGCTGAGAGAGCCTGGGCAAGAGACAGAGAGG 138
DB 127 CAGGCCCAATATGTGGCCCTATCGCTGAGAGAGCCTGGGCAAGAGACAGAGAGG 186
QY 139 CTGCTTAACCGGGGTCACATGATCAGACGAGTGGTGGCTCAAGATCGGAAC 198
DB 187 CTGCTTAACCGGGGTCACATGATCAGACGAGTGGTGGCTCAAGATCGGAAC 246
QY 199 CGGAGAGAGTGTGGAGTGTGGTGTGATGAGAGTGGAGAGTCCCATCTGAAG 258
DB 247 CGGAGAGAGTGTGGAGTGTGGTGTGATGAGAGTGGAGAGTCCCATCTGAAG 306
QY 259 CTGATCGAAGCCCAATGCTCTCAAGCTCAAGAGCTTACAGAGAAAGAAATTTG 318
DB 307 CTGATCGAAGCCCAATGCTCTCAAGCTCTCAAGAGCTTACAGAGAAAGAAATTTG 366
QY 319 TACCTGGTTCGAGCAGTCTCGGGGGTGGATTCGATCTACTGGTAAAGAGGG 378
DB 367 TACCTGGTTCGAGCAGTCTCGGGGGTGGATTCGATCTACTGGTAAAGAGGG 426
QY 379 AGACTGAGCCCAAGAGAGCCCGAAGTTCTCCGCAAGATTTGTCCTGCTGACTTC 438
DB 427 AGACTGAGCCCAAGAGAGCCCGAAGTTCTCCGCAAGATTTGTCCTGCTGACTTC 486
QY 439 TGCACAGACTTACTCATCTGACAGAGACTTAAAGCCGGAAGACTGCTTTGATGAG 498
DB 487 TGCACAGACTTACTCATCTGACAGAGACTTAAAGCCGGAAGACTGCTTTGATGAG 546
QY 499 AAAAAACAATCCGATTCGAGACTTGGCATGGGCTGCTCGAGTGGGGGACACCTTC 558
DB 547 AAAAAACAATCCGATTCGAGACTTGGCATGGGCTGCTCGAGTGGGGGACACCTTC 606
QY 559 CTGAGAGACAGTGGGGTCCCTCCATTTATGCTGTCCAGAGTGAATTAAGGGGAAAAA 618
DB 607 CTGAGAGACAGTGGGGTCCCTCCATTTATGCTGTCCAGAGTGAATTAAGGGGAAAAA 666
QY 619 TATATGAGCCCGGGGAGACATGTGAGCTGTGAGTCACTCTTCTGCTGCTGCTG 678
DB 667 TATATGAGCCCGGGGAGACATGTGAGCTGTGAGTCACTCTTCTGCTGCTGCTG 726
QY 679 GGGGCTCTGCTTGTGAGAGACAACTCCGCGAGCTGCTGAGAAAGTGAACGGGGG 738

DB 727 GGGGCTCTGCTTGTGATGACAGCAACTCCGCAAGCTGCTGAGAGAGTGAACGGGGG 786
QY 739 GTCTTCACATGCCCCCACTTCTTCATTCCTCAGATTGCCAGAGGCTCTTGAGGAATGATC 798
DB 787 GTCTTCACATGCCCCCACTTCTTCATTCCTCAGATTGCCAGAGGCTCTTGAGGAATGATC 846
QY 799 GAACTGAGCCCGAAAAAGGCTCAGTCTGAGCAAAATTCAGAAACATCTTGTGATCTA 858
DB 847 GAACTGAGCCCGAAAAAGGCTCAGTCTGAGCAAAATTCAGAAACATCTTGTGATCTA 906
QY 859 GGGGGGAAACAGAGCCAGACCCCGCTGAGAGCCAGCCCGCTGGGCTGAGGATGATC 918
DB 907 GGGGGGAAACAGAGCCAGACCCCGCTGAGAGCCAGCCCGCTGGGCTGAGGATGATC 966
QY 919 CGGAGCTGCTCATCAACAGAGAGTGGACCCGAGCTCTGAGAGAGTGCATCACTG 978
DB 967 CGGAGCTGCTCATCAACAGAGAGTGGACCCGAGCTCTGAGAGAGTGCATCACTG 1026
QY 979 GGGTCTTCAGGAGCCGAGAGAGCTGCATGCGAGCTGCGCAGTGAAGAGAGAACCA 1038
DB 1027 GGGTCTTCAGGAGCCGAGAGAGCTGCATGCGAGCTGCGCAGTGAAGAGAGAACCA 1086
QY 1039 GAAAGAGAT 1098
DB 1087 GAAAGAGAT 1146
QY 1099 CAGGACCTGCTCCCGGAATGATGTGACCCCGCCGGAAGCTGTGATCTCCATG 1158
DB 1147 CAGGACCTGCTCCCGGAATGATGTGACCCCGCCGGAAGCTGTGATCTCCATG 1206
QY 1159 CTGAGCGTCAAGGAGAGCCGAGACAGAGCGGAAGTCAATGAAGTCTGAGATCAAC 1218
DB 1207 CTGAGCGTCAAGGAGAGCCGAGACAGAGCGGAAGTCAATGAAGTCTGAGATCAAC 1266
QY 1219 GATGCGGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1278
DB 1267 GATGCGGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1326
QY 1279 AGCCAGAGATCCCTAGAGTGTGAGAGCTTCACAGGCTGTCTCCAGCCCTTAAGC 1338
DB 1327 AGCCAGAGATCCCTAGAGTGTGAGAGCTTCACAGGCTGTCTCCAGCCCTTAAGC 1386
QY 1339 AGCCAAAGAGTCCGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1398
DB 1387 AGCCAAAGAGTCCGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1446
QY 1399 GGGGGGGGCTCCCGAGTCTCAAAAGCAGAGAGCTGCTTCTCGGGGGGCCCCAGGGGTGG 1458
DB 1447 GGGGGGGGCTCCCGAGTCTCAAAAGCAGAGAGCTGCTTCTCGGGGGGCCCCAGGGGTGG 1506
QY 1459 GGGGCGGGGAGAGCCCGCCCGCCAGTGCAGCTCCGCTCCACACCCCTCCCGGGCCCCCA 1518
DB 1507 GGGGCGGGGAGAGCCCGCCCGCCAGTGCAGCTCCGCTCCACACCCCTCCCGGGCCCCCA 1566
QY 1519 GGGTCCCGGGCTCTCTGCGGAGCCCTTTGCACTGCTCTGACAGCCCGGGGCT 1578
DB 1567 GGGTCCCGGGCTCTCTGCGGAGCCCTTTGCACTGCTCTGACAGCCCGGGGCT 1626
QY 1579 AGTCCACCCGGGAGCCCGGGGAGCAACACACCCCGGAGCGGGGGGTGGGGGA 1638
DB 1627 AGTCCACCCGGGAGCCCGGGGAGCAACACACCCCGGAGCGGGGGGTGGGGGA 1686
QY 1639 GCGGCTGAGAGAGTGTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1698
DB 1687 GCGGCTGAGAGAGTGTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1746
QY 1699 CAGCGGCGAAGATGAGAGTCCCTACCGCTGAGAGATGTCAGCTTGAACGCGAGATCC 1758
DB 1747 CAGCGGCGAAGATGAGAGTCCCTACCGCTGAGAGATGTCAGCTTGAACGCGAGATCC 1806
QY 1759 TCCCGGAGCTGGCAAAAGCTCTGCTGTGGGAACTTCACTCTCTTGGACAAAGAGAA 1818
DB 1807 TCCCGGAGCTGGCAAAAGCTCTGCTGTGGGAACTTCACTCTCTTGGACAAAGAGAA 1866

PR	15-JUN-2000;	2000US-212073P.
PR	23-JUN-2000;	2000US-213467P.
PR	30-JUN-2000;	2000US-215651P.
PR	07-JUL-2000;	2000US-216605P.
PR	13-JUL-2000;	2000US-218372P.
PR	25-AUG-2000;	2000US-228056P.
XX		
XX	(INCY-)	INCYTE GENOMICS INC.
XX		
XX	Yue H, Lal P, Bandman O, Borowsky ML, Au-Young J, Lu Y;	
XX	Gandhi AR, Tribouley CM, Wallia NK, Yao MG, Lu DM, Greenwald SR;	
XX	Ramkumar J, Griffin JN, Kearney L, Burford N, Nguyen DB, Tang YT;	
XX	Baughn MR, He A, Thornton M, Hafalia A, Patterson C, Gururajan R;	
XX	Lo TP, Khan F, Reippon SA, Azimzai Y, Policky JL, Ding L,	
XX	Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;	
XX	WPI: 2002-090207/12.	
XX	P-PSDB: AAE16271.	
PT	New polypeptides, useful for diagnosing, treating or preventing	
PT	disorders of growth and development, cardiovascular and lipid, and	
PT	diseases such as cancer, comprise human kinase polypeptides -	
PS	Claim 5; Page 188-189; 197pp; English.	
XX		
CC	The invention relates to human kinase PKIN proteins and their	
CC	corresponding cDNAs. A composition containing PKIN agonist is useful for	
CC	treating a disease or condition associated with decreased expression of	
CC	PKIN and a composition comprising PKIN antagonist is useful for treating	
CC	a disease or condition associated with overexpression of PKIN. The	
CC	disorders include cancer (leukemia, adenocarcinoma, lymphoma, melanoma,	
CC	myeloma, sarcoma, teratocarcinoma, Hodgkin's disease); immune disorder	
CC	(Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease,	
CC	atherosclerosis, anemia, allergies, adult respiratory distress syndrome,	
CC	autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes	
CC	mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease,	
CC	osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome,	
CC	rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis,	
CC	bacterial, parasitic, fungal, viral, protozoal and helminthic infections)	
CC	growth and development disorders (arteriosclerosis, cirrhosis, hepatitis,	
CC	Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts); cardio	
CC	vascular disease (arteriovenous fistula, hypertension, vasculitis,	
CC	aneurysms, congestive heart failure, angina pectoris, myocarditis,	
CC	ischemic heart disease, chronic bronchitis, lung tumours); lipid	
CC	disorder (fatty liver, Fabry's disease, Niemann-Pick's disease,	
CC	hypercholesterolemia, obesity). PKIN DNA is useful for assessing	
CC	toxicity of a test compound and in gene therapy.. The present sequence	
CC	is human PKIN-17 cDNA.	
XX		
XX	Sequence 2897 BP; 599 A; 921 C; 877 G; 500 T; 0 other;	
QQ		
QQ	Query Match 96.7%; Score 2256; DB 24; Length 2897;	
QQ	Best Local Similarity 100.0%; Pred. No. 0;	
QQ	Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	79 CACGCCCAATATGTGGGCCCTATCGCTGAGAGACGCTGGGCAAGAGCAGACAGGG 138	
Db	127 CACGCCCAATATGTGGGCCCTATCGCTGAGAGAGCAGCTGGGCAAGAGCAGACAGGG 186	
QY	139 CTGGTTAAACTGGGGGTCCATCTGCATCAACGGGTCAAGAGTGGCCATCAAGATCGTGAAC 198	
Db	187 CTGGTTAAACTGGGGGTCCATCTGCATCAACGGGTCAAGAGTGGCCATCAAGATCGTGAAC 246	
QY	199 CGGGAAGAGCTGTGGAGCTGGTGTGTGATGAGAGTGGAGCGGGAAGATGCCATCCTGAAG 258	
Db	247 CGGGAAGAGCTGTGGAGCTGGTGTGTGATGAGAGTGGAGCGGGAAGATGCCATCCTGAAG 306	
QY	259 CTCATCGAAGACCCACATGTCTCTCAAGCTCCACGACGCTCTACGAAACAGAAATATTGG 318	
Db	307 CTCATCGAAGACCCACATGTCTCTCAAGCTCCACGACGCTCTACGAAACAGAAATATTGG 366	
QY	319 TACCTGTTCTTGAGACACGCTCTGGGGGGGTGAGCTATTTCAGCTACCTGGTTAAAGAGGGG 378	

Db 367 TACCTGGTTGAGACACGCTCGGGGGGTGAGTATTCGACTACCTGGTAAAGAGG 426
 QY 379 AGACTGAGCCCAAGAGAGCCCGAAGTTCTTCCGACAGATTGTGTCGCTGAGACTTC 438
 Db 427 AGACTGAGCCCAAGAGAGCCCGAAGTTCTTCCGACAGATTGTGTCGCTGAGACTTC 486
 QY 439 TGGCAGACGTAATCTGATGTCACAGAGACCTAAAGCCCGAGAACCTGCTTTTGATGAG 498
 Db 487 TGGCAGACGTAATCTGATGTCACAGAGACCTAAAGCCCGAGAACCTGCTTTTGATGAG 546
 QY 499 AAAAACAACATCCGATGAGACGTAATGCGCATGGGTCCCTGACAGTGGGGGACACCTTC 558
 Db 547 AAAAACAACATCCGATGAGACGTAATGCGCATGGGTCCCTGACAGTGGGGGACACCTTC 606
 QY 559 CTGAGAGACGTAATGAGGTCCTCCCATTTATGCGTGTCCAGAGTGTATTAAGGGGAAAA 618
 Db 607 CTGAGAGACGTAATGAGGTCCTCCCATTTATGCGTGTCCAGAGTGTATTAAGGGGAAAA 666
 QY 619 TATGATGGCCCGCGGAGACATGTGAGCTGTGAGTCAATCCTCTTCGCTGCTGATG 678
 Db 667 TATGATGGCCCGCGGAGACATGTGAGCTGTGAGTCAATCCTCTTCGCTGCTGATG 726
 QY 679 GGGGCTCGCCCTTTGATGAGACAACTCCGCGCAGCTGCTGGAAGAGTGAACGGGGG 738
 Db 727 GGGGCTCGCCCTTTGATGAGACAACTCCGCGCAGCTGCTGGAAGAGTGAACGGGGG 786
 QY 739 GTCTTCACATGCCCCACTTATTCCTCAGATTGCCAGAGCTCCTTGAGGGGAATGATC 798
 Db 787 GTCTTCACATGCCCCACTTATTCCTCAGATTGCCAGAGCTCCTTGAGGGGAATGATC 846
 QY 799 GAACTGGAGCCGAAAAAGGCTCAGTCTGAGCAAAATTCAGAAACATCTTGTGACTA 858
 Db 847 GAACTGGAGCCGAAAAAGGCTCAGTCTGAGCAAAATTCAGAAACATCTTGTGACTA 906
 QY 859 GGGGGAAGAAAGAGAGCCGAGACCCGTCGAGAGCCGCGGGGGGAGGAGCATG 918
 Db 907 GGGGGAAGAAAGAGAGCCGAGACCCGTCGAGAGCCGCGGGGGGAGGAGCATG 966
 QY 919 CGAGAGCTGCCATCCAAAGAGAGCTGAGACCCGAGCTCTAGAGAGCATGGCATTCTG 978
 Db 967 CGAGAGCTGCCATCCAAAGAGAGCTGAGACCCGAGCTCTAGAGAGCATGGCATTCTG 1026
 QY 979 GGGCTCTTCAGAGAGCCGAGAGAGTGCATGGCAGCTGCGAGTGAAGAGAGAGAAACCA 1038
 Db 1027 GGGCTCTTCAGAGAGCCGAGAGAGTGCATGGCAGCTGCGAGTGAAGAGAGAGAAACCA 1086
 QY 1039 GAAAGATGATATATATATCTGCTTTGGATGGAAGAGAGCGGTATCCCGAGCTGTGAGAC 1098
 Db 1087 GAAAGATGATATATATATCTGCTTTGGATGGAAGAGAGCGGTATCCCGAGCTGTGAGAC 1146
 QY 1099 CAGGAGCTGCCCTCCCGAATGATGTGACCCCGCGAAGCGTGTGATTCCTCCATG 1158
 Db 1147 CAGGAGCTGCCCTCCCGAATGATGTGACCCCGCGAAGCGTGTGATTCCTCCATG 1206
 QY 1159 CTGAGCCGTCACGGGAAGCGCGAGACAGAGGGAATTCATGGAAGTCTGTGAGCATCAC 1218
 Db 1207 CTGAGCCGTCACGGGAAGCGCGAGACAGAGGGAATTCATGGAAGTCTGTGAGCATCAC 1266
 QY 1219 GATGCGGGGGTGTGGTCCCTGTAACCAACCGAGGGGCTTGGAGATGGCCCGAGAC 1278
 Db 1267 GATGCGGGGGTGTGGTCCCTGTAACCAACCGAGGGGCTTGGAGATGGCCCGAGAC 1326
 QY 1279 AGCCAGAGATCCCTAGACGTCAGTGAGACCTCCACGGGTGTGTCTCTCCAGCCCTCTAAC 1338
 Db 1327 AGCCAGAGATCCCTAGACGTCAGTGAGACCTCCACGGGTGTGTCTCTCCAGCCCTCTAAC 1386
 QY 1339 AGCCAGAGATCCCTAGACGTCAGTGAGACCTCCACGGGTGTGTCTCTCCAGCCCTCTAAC 1398
 Db 1387 AGCCAGAGATCCCTAGACGTCAGTGAGACCTCCACGGGTGTGTCTCTCCAGCCCTCTAAC 1446
 QY 1399 GGGGGGGGCTCCCGAGACTTCCAAAAAGCAGAGCTGCTTCCGGGGGGGGGGGGGGGG 1458
 Db 1447 GGGGGGGGCTCCCGAGACTTCCAAAAAGCAGAGCTGCTTCCGGGGGGGGGGGGGGGG 1506

QY 1459 GGGCGGGGAGACCCCGCCCGCCAGTGGCGCTCCACACCCCTGCGGGGGGGGG 1518
 Db 1507 GGGCGGGGAGACCCCGCCCGCCAGTGGCGCTCCACACCCCTGCGGGGGGGGG 1566
 QY 1519 GGGTCCCGGGCTCTCTGCGGGAGCCCGCTTGACACTGCGCTCTGACAGCGCCGGGG 1578
 Db 1567 GGGTCCCGGGCTCTCTGCGGGAGCCCGCTTGACACTGCGCTCTGACAGCGCCGGGG 1626
 QY 1579 AGTCCACCGGGAGCCCGGGGACAAACACACCCCGCCAGCCCGGGGGGGG 1638
 Db 1627 AGTCCACCGGGAGCCCGGGGACAAACACACCCCGCCAGCCCGGGGGGGG 1686
 QY 1639 GCGGCTGAGAGAGTCTGCTCACTCCATCCGCAACACCTTCGCGGGTCCCGCTGCTT 1698
 Db 1687 GCGGCTGAGAGAGTCTGCTCACTCCATCCGCAACACCTTCGCGGGTCCCGCTGCTT 1746
 QY 1699 CACCGGGCAAGATGACAGTCCCTACCGCTGAGAGATGTCCAGCTTACGCGAGATCC 1758
 Db 1747 CACCGGGCAAGATGACAGTCCCTACCGCTGAGAGATGTCCAGCTTACGCGAGATCC 1806
 QY 1759 TCCCGGAGCTGGGCAAAAGCTCTGCTGCTGCGGAATTCATCTCTTGGACAAAGAA 1818
 Db 1807 TCCCGGAGCTGGGCAAAAGCTCTGCTGCTGCGGAATTCATCTCTTGGACAAAGAA 1866
 QY 1819 CAAATATTCCTGCTGCTAAAGCAAACTCTCAGCAGATCAAAAGCAATGCTCAT 1878
 Db 1867 CAAATATTCCTGCTGCTAAAGCAAACTCTCAGCAGATCAAAAGCAATGCTCAT 1926
 QY 1879 GCGTCTGTGATGATCCCGAGCTGAGTCAAGTGTCTGTACAGACAGCTTCAGGGCC 1938
 Db 1927 GCGTCTGTGATGATCCCGAGCTGAGTCAAGTGTCTGTACAGACAGCTTCAGGGCC 1986
 QY 1939 GAGTACAGGCGAGTGGCGGGCGCTCCGCTCTCCAAAGCCCGGCTTCAGGTGAGC 1998
 Db 1987 GAGTACAGGCGAGTGGCGGGCGCTCCGCTCTCCAAAGCCCGGCTTCAGGTGAGC 2046
 QY 1999 ATCAGCTCTCTGAGGGTCCAGAGCCCTCCCGGAGCGAGCGAGGAGTGTGGC 2058
 Db 2047 ATCAGCTCTCTGAGGGTCCAGAGCCCTCCCGGAGCGAGCGAGGAGTGTGGC 2106
 QY 2059 ATCTACTCCGTCACCTTCACTCTCATCTCGGGTCCAGCCGTCGGTTCAGGAGTGTG 2118
 Db 2107 ATCTACTCCGTCACCTTCACTCTCATCTCGGGTCCAGCCGTCGGTTCAGGAGTGTG 2166
 QY 2119 GAGACATCCAGGACACAGCTCTGAGACATGACAGCCCTCCGTCAGGAGCCCTGGA 2178
 Db 2167 GAGACATCCAGGACACAGCTCTGAGACATGACAGCCCTCCGTCAGGAGCCCTGGA 2226
 QY 2179 GAGGAGAAAGCGGGGCGCAGACCCGGCTGCTGCTGCTCCACCCGAAAGCTGAGGCC 2238
 Db 2227 GAGGAGAAAGCGGGGCGCAGACCCGGCTGCTGCTGCTCCACCCGAAAGCTGAGGCC 2286
 QY 2239 CCACCGGGGCGCCAGACCCAGAGCTGAGACGCTCTCCCGCGAGGCCCGCCAGAGAC 2298
 Db 2287 CCACCGGGGCGCCAGACCCAGAGCTGAGACGCTCTCCCGCGAGGCCCGCCAGAGAC 2346
 QY 2299 AAGAAGCTCTGGGCAACCAAGGAGCCCTGAGCC 2354
 Db 2347 AAGAAGCTCTGGGCAACCAAGGAGCCCTGAGCC 2382

RESULT 3
 ABA02995
 ID ABA02995 standard; cDNA; 2025 BP.
 XX ABA02995;
 AC
 XX
 DT 19-FEB-2002 (first entry)
 XX
 DE Human protein kinase 2246 coding sequence SEQ ID NO. 3.
 XX
 KW Human; protein kinase 2246; cytosolic; immunomodulator; carcinoma;

KW anti-inflammatory; analgesic; cardiovascular; cancer; sarcoma;
 KW cellular proliferation disorder; cellular differentiation disorder;
 KW metastatic; haematopoietic disorder; leukaemia; immune disorder;
 KW inflammatory disorder; arthritis; autoimmune disease; diabetes mellitus;
 KW psoriasis; Crohn's disease; cardiovascular disease; virus; pain;
 KW gene therapy; ss.
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT 1..2025
 FT CDS /tag- a
 FT /product- "protein kinase 2246"
 PN MO200181588-A2.
 XX 01-NOV-2001.
 XX 25-APR-2001; 2001MO-US13784.
 XX 25-APR-2000; 2000US-199391P.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX
 XX
 PI Meyers R;
 DR WPI: 2002-049281/06.
 DR P-PSDB; AAA47830.
 XX
 PT New protein kinase nucleic acid and polypeptide molecules, designated
 PT 2246, useful for diagnosing, preventing or treating cancer or a
 PT cellular proliferation/differentiation disorders, e.g. carcinoma,
 PT sarcoma or leukaemias -
 XX
 PS Claim 1; Fig 1; 11pp; English.
 XX
 CC The invention relates to the human protein kinase 2246 gene and the
 CC the isolated encoded polypeptide with cytosolic, immunomodulator,
 CC anti-inflammatory, analgesic and cardiovascular activity. The 2246
 CC nucleic acid and polypeptide are useful for diagnosing, preventing or
 CC treating a subject having cancer or a cellular proliferation and/or
 CC differentiation disorder or at risk of developing cancer or a cellular
 CC proliferation and/or differentiation disorder. In particular, the
 CC disorder includes carcinoma, sarcoma, metastatic or haematopoietic
 CC disorders (e.g. leukaemias) or cancers of the lung, breast, thyroid, head
 CC neck, prostate or genito-urinary tract. The 2246 nucleic acid and
 CC polypeptide are also useful for treating immune disorders, e.g.
 CC inflammatory (e.g. respiratory inflammation or arthritis), autoimmune
 CC disease (e.g. diabetes mellitus, psoriasis, Wegener's granulomatosis,
 CC Crohn's disease or Grave's disease), for treating cardiovascular
 CC diseases, endothelial cell disorder, viral diseases or pain. The nucleic
 CC acid and polypeptide are also useful for evaluating the efficacy of a
 CC treatment of cancer or a cellular proliferation and/or differentiation
 CC disorder. The nucleic acid is useful for gene therapy. The present
 CC sequence is that of the 2246 coding sequence.
 CC
 CC
 XX
 SQ Sequence 2025 BP; 442 A; 641 C; 616 G; 326 T; 0 other;
 Query Match 41.28; Score 961; DB 24; Length 2025;
 Best Local Similarity 70.18; Pred. No. 1.1e-199;
 Matches 1477; Conservative 0; Mismatches 435; Indels 195; Gaps 6;

254 TGAAGTCATCGAACACCCACATGTCCTCAAGTCACGACGTGTAGAGAACAAAT 313
 209 TGAAGTCATGAGACACCCACAGTCTAAAGTCACGACGAGTGTAGAAACAAAT 268
 314 ATTGTACCTGCTGTGAGACAGCTGTGGGGGGGTAGCTATTGCACTACCTGTAAAG 373
 269 ATTGTACCTGCTGTGAGACAGCTGTGGGGGGGTAGCTATTGCACTACCTGTAAAG 328
 374 AGGGGAGCTGAGCCCAAGAGGGCCGAAAGTCTCCGCAATGTGTCTGCGCGG 433
 329 AGGGGAGCTGAGCCCAAGAGGGCCGAAAGTCTCCGCAATGTGTCTGCGCGG 388
 434 ACTTTCGCCACACTTCCATCTGCCACAGAGACTAAAGCCGAGAACCTGTGTTGG 493
 389 ACTTTCGCCACACTTCCATCTGCCACAGAGACTAAAGCCGAGAACCTGTGTTGG 448
 494 ATGAGAAAACAAACATCCGATTTGACACTTGGCATGGCTCCCTGAGGTGGGGACA 553
 449 ACAGAAAGAAACAAACATCCGATTTGACACTTGGCATGGCTCCCTGAGGTGGGGACA 508
 554 GCCTCTGAGACACAGTGGGGGTCCCCCATTTAGGTGTCCAGAGGTGATTAAGGGG 613
 509 GCCTGTGAGACACAGTGGGGGTCCCCCATTTAGGTGTCCAGAGGTGATTAAGGGG 568
 614 AAAAATATGATGGCCGCGGAGACATGTGACCTGTGAGTATCTTCCGCTGC 673
 569 AGAATATGATGGCCGCGGAGACATGTGACCTGTGAGTATCTTCCGCTGC 628
 674 TCGTGGGGCTGCTCCCTTTGATGACGACAACTCCGACCTGTGTGAGAGTGAAC 733
 629 TCGTGGGGCTGCTCCCTTTGATGACGACAACTCCGACCTGTGTGAGAGTGAAC 688
 734 GGGGCGCTTCCACATCCCATCTTCAATCTCCAGATGCGAGCTCTGAGAGGAA 793
 689 GGGGCGCTTCCACATCCCATCTTCAATCTCCAGATGCGAGCTCTGAGAGGAA 748
 794 TGTGCAAGTGAAGCCGAAAAAGGCTCACTGTGAGCAAAATTCAGAACATCTTGT 853
 749 TGAAGCAGGTGAGACGCGGACGCGCTCAGCTAGACATTCAGAACATATGATGT 808
 854 ACCTAGCGGGGAAACAGACGACGACCGCTGTGAGACCGCTGTGCGGGGGGTAG 913
 809 ATATAGGGGCAAAATGAGCCGCAAC-----AGACCAACCCATCTTCCGCAAGTGC 862
 914 CCAATGCGAGCTGCTCCATCCAAAGAGTGAAGCCGAGCTGCTTAGAGACATGAT 973
 863 AGATCCGCTGCTGCTCCCAAGCTGAGAGCATGACCCGAGCTGTGACAGACATGCT 922
 974 CACTGGCTCTTCAAGGACCGGAGAGCTGATCCGAGCTGTGAGAGAGAGA 1033
 923 CACTGGCTCTTCAAGGACCGGAGAGCTGATCCGAGCTGTGAGAGAGAGA 982
 1034 ACCAAGAAAAGATATATATATCTGCTTTGATGAGAAAGAGGATATCCACTGTG 1093
 983 ACCAAGAAAAGATATATATATCTGCTTTGATGAGAAAGAGGATATCCACTGTG 1042
 1094 AGGACAGAGACTGCTCCCGGATGATGTTGACCCCGGAGAGGTGTGATGCTC 1153
 1043 AGGATGAGAGACTGCTCCCGGATGATGTTGACCCCGGAGAGGTGTGATGCTC 1102
 1103 CGATGCTGAACCGGACGAGGCAAGCGGCGGACAGAAAGCAATTCATGAGAGGTCTCAGCG 1162
 1214 TCACCGATGCGCGGGGTGTGCTGCTTGTATCCACCCGAGCGGCTTGTGAGATGGCC 1273
 1163 TGAC-----GAGCGCGGCTGCTGCTTGTATCCACCCGAGCGGCTTGTGAGATGGCC 1213
 1274 AGCAGCAGAGAGATCCCGTATGCTAGTGAAGCTCCACAGGCTGTCTTCCAGCCCTC 1333
 1214 AGCAGCAGAGAGATCCCGTATGCTAGTGAAGCTCCACAGGCTGTCTTCCAGCCCTC 1273

[illegible]

KW cellular proliferation disorder; cellular differentiation disorder;
KW metastatic; haematopoietic disorder; leukaemia; immune disorder;
KW inflammatory disorder; arthritis; autoimmune disease; diabetes mellitus
KW psoriasis; Crohn's disease; cardiovascular disease; virus; pain;
KW gene therapy; ss.
XX
XX Homo sapiens.
OS
XY

Key CDS	Location/Qualifiers
	51..2075
	/*tag= a
	/product= "protein kinase 2246"

MO200181588-A2

01-NOV-2001

25-APR-2001; 2001WO-US13784.

25-APR-2000; 2000US-199391P.

(MILL-) MILLENNIUM PHARM INC.

Meyers R,

WPI: 2002-049281/06.

P-PSDB; AAM47830.

New protein kinase nucleic acid and polypeptide molecules, designated

cellular proliferation/differentiation disorders, e.g. carcinoma,

Claim 1; Fig 1; 111pp; English.

The invention relates to the human protein kinase 2246 gene and the isolated encoded polypeptide with cytoskeletal, immunomodulatory, anti-inflammatory, analgesic and cardiovascular activity. The 2246 nucleic acid and polypeptide are useful for diagnosing, preventing or treating a subject having cancer or a cellular proliferation and/or differentiation disorder or at risk of developing cancer or a cellular proliferation and/or differentiation disorder. In particular, the disorder includes carcinoma, sarcoma, metastatic or haematopoietic disorders (e.g. leukemias) or cancers of the lung, breast, thyroid, neck, prostate or genito-urinary tract. The 2246 nucleic acid and polypeptide are also useful for treating immune disorders, e.g. inflammatory (e.g. respiratory inflammation or arthritis), autoimmune disease (e.g. diabetes mellitus, psoriasis, Wegener's granulomatosis, Crohn's disease or Grave's disease), for treating cardiovascular diseases, endothelial cell disorder, viral diseases or pain. The nucleic acid and polypeptide are also useful for evaluating the efficacy of a treatment of cancer or a cellular proliferation and/or differentiation disorder. The nucleic acid is useful for gene therapy. The present sequence is that of the 2246 encoding cDNA.

Sequence 2217 BP; 476 A; 715 C; 682 G; 344 T; 0 other;

Query Match	41.28;	Score 961;	DB 24;	Length 2217;

Matches 1477; Conservative 0; Mismatches 435; Indels 195; Gaps 6;

74 CCAGCAGCCCAATATGTGGCCCTATCGCTGGAGAGACGCTGGCCAAGGACAGA 133

Db 79 CGCAGCACGGCAGTATGTTGGCCCTACCGGCTGGAGAGACGCTGGGCAAGGGCAGA 138

QY 134 CAGGGCTGGTTAACTCGGGGTCCTGCATCACGGGTCAGAAGTCCCATCAAGATCG 193

Db 139 CAGGTCGTGAAGCTGGGGGTTCACTGCCGTCACTGCCAGAGGTGGCCATCAGATCG 198

194 TGACCGGAGAGCTGTCGGAGTCGGTCTGATGAAGGTGGAGCGGAGATCGCCATCC 253

Db 199 TCACCGTGAGAGCTCAGCGAGTCCGCTGATGAAGGTGGAGCGGAGATCCGATCC 258

QY	254	TGAACCTCATGCAGAACCCACACATATGCTCCATAGCTCCAGACACAGCTCTACGAAACAAGAAAT	313
Db	259	TGAAGCTCATTTAGAGACACCCCCACGTCCTTAAAGCTGCACGACGTTTATGAAAAACAAAAAT	318
QY	314	ATTGTACCTGGTTTGTGAGCACGTCCTCGGGGGGTAGAGCTATTGCACTACCTGGTAAAGA	373
Db	319	ATTGTACTGTGTGTAGAAACACAGTGTACAGTGTGTAGAGCTCTTCGACTACCTGTGTAAAGA	378
QY	374	AGGGGACACTGACGCCCAAGAGAGGCCCAAGATTCTTCCGCCACGATTTGTCTTCGCTGG	433
Db	379	AGGGAGAGCTGACGCTTAAGGAGAGGCTCGGAAGTTCTTCCGCGACGATCATCTCTGCGCTGG	438
QY	434	ACTTTGGCACAGCTACCTCATCTGGCCACAGAGACCTTAAGGCCGGAACCTGCTTTGG	493
Db	439	ACTTTGGCCACGCCCATCTCATATGCTCCACAGGGATTTGAAACCTGTAAACCTCTGCTGG	498
QY	494	ATGAGAAAAACACATCCGATTTGCGAGCTTGGCGATGGCGTCTCTGACAGTGGGGACACA	553
Db	499	ACGAAAGACACAACTCCGCATTCGGAGATTTGGCATTTGGGTCTCTGACAGTTGGCGACA	558
QY	554	GCCCTCTGGAGACACAGCTGCGGGGTCCCCCATTTATCGTGTCCAGAGTGATTAAAGGGG	613
Db	559	GCCCTTTGGAGACACAGCTGTGGGTCCCCCACTACGCTGCCCCGAGAGTATCCGGGGGG	618
QY	614	AAAAATATGATGTCGGCCGGCGGACAGCATGTGGAGCTGTGGAGTCATCCTCTTGGCCCTGC	673
Db	619	AGAAATATGATGACGGCCGGAGAGGGAGGTGTGGAGCTGGCGCGTCACTCTGTGCTTCGCTTGC	678
QY	674	TGCTGGGGGCTTGGCCCTTTGATGAGCAACACTCCGCCACGTGCTGAGAAAGGTGAAC	733
Db	679	TGCTGGGGGCTTGGCCCTTCGAGATGACACATTTGGGACAGCTGCTGGAAAGGTGAAC	738
QY	734	GGGGGCTCTTCCACATGCCCCCACTTCATTCCTCCAGATTGCCAGAGCTCTCTGAGGGAA	793
Db	739	GGGGGCTGTTCCACATGCGCGCACTTATCCCGCCCACTGCCAGAGCTCTCTAGGGGCA	798
QY	794	TGATGAGTGTGAGGCCGCGGAAAAAGGCTCAGTGTGGAGCAAAATTCGAAACATCCTTGGT	853
Db	799	TGAGGAGGTGAGACGGCCGACGCGGCTCACTGCTAAGACACATTCAAAAAACATATAGT	858
QY	854	ACCTTAGGCGGAAACACGAGCCAGACCCGTGCTGTGAGACGAGCCCGCGCGCGGGTAG	913
Db	859	ATATAGGGGGGACGAAGATGAGCCGAAAC-----AAGACGCCCATTTCTGCGAAGGTGC	912
QY	914	CCATGCGGAGCCTGCTCCATCCAAACGAGAGCTGAGACCCCGACGTCCTTAGAGACATGGCAT	973
Db	913	AGATCCGCTCGCTGCCACCTGTGAGAGCAATCGAACCCCGACGCTGTGAGACATGACACT	972
QY	974	CACATGGGCTGCTTACAGGAGACCGCGAGAGGCTGCATGCGCAGCTGCGCATGTAGAGAGAGA	1033
Db	973	CACATGGGCTGCTTCCGAGACCGCCACAACTGCTGTGAGAGCACTGTGCTCCGAGAGAGAGA	1032
QY	1034	ACCAAGAAAAAGATATATATATATCTCTTTTGATATGGAAGGAGCGGTATGCCAGCTGTG	1093
Db	1033	ACCAAGAGAAAGATGATTTTCTTCTCTCTCTGACCGGAAAGAAAGGTATCCCGAGCGAGG	1092
QY	1094	AGGACAGGACCTGCTCCCGGAAATGATGTGACCCCGCCCGGAAGCGTGTGATTTCTC	1153
Db	1093	AGGATGAGGAACTGCCCCCGGAAAGAGATGATGACCTCCCGGAAGCGTGTGAACTGCC	1152
QY	1154	CCATGCTGAGCCGTACAGGGAAAGCGCGACACAGAGGAAAGTCCATGAAAGTCTTGACA	1213
Db	1153	CGATCTGTAACCGGACCGGCAAGCGCGGCGCAACAGCAATCATGATGAGTGTCTCAGCG	1212
QY	1214	TCACGAGTCCCGGGGTGGTGGCTCCCGTGTACCAACCGACGGGCGCTTGGAGATGGGCC	1273
Db	1213	TCAC-----GACGGCGGCTCCCGGCTGCTGCGCGCGGCGCATTTGAGATGGGCC	1253
QY	1274	AGCAGCGCCAGAGATCCCTAGCGCTCAGTGAAGCCTTCCACGGGATCTGCTCTCCACGCCCTC	1333
Db	1264	ACGACGGCCAGAGGTCTGTGGTTCATCAGGCGGTGCTCTCAGGCGCTTTCACACAGGCCAC	1323
QY	1334	TAAAGCAGCCAAAGAGTCCGGTCTTTTCTTTTCCACCGGAGCCGGGGCTGGAGATGAGG	1393

Db	1324	TCACGACGCCCCG	-----	1336
QY	1394	CTGAGGCGGGGGCTCCCCGACTTCCAAAAAGCAGACGTCGCTTCTCGGGGCCCCAGGG		1453
Db	1337	-----	-----	1336
QY	1454	GTTGGGGGCGCGGGGAGCAGCCCCCCCCCACTGCGCGCTCCACACCCCTGCGGGCC		1513
Db	1337	-----	-----	1376
QY	1514	CCCCAGGCTCCCGCGCTCCCTCTGCGGGGAGACCCCTTGACACTGCGCTCTGACACAGCCCC		1573
Db	1377	-----	-----	1407
QY	1574	GGGCGAGTCCACCCGGGAGACCCCGGGGCAACACACACCCCCAGCCCCCGCGGTGGCTCG		1633
Db	1408	AGGAGAGCGCGGGCTGGCACGCGCCACCCAGCGCCCCGTCGACGCCC	-----	1461
QY	1634	GGGAGCGCGCTGGAAGAGTGTCTCACTCATCCGCAACAGCTTCTGCGCTCCCTC		1693
Db	1462	GAGGGGTGCCCTTGAGGGGCGCGGCTCAACTCAATCAAGAAACAGTTTCTGGGCTCACCCC		1521
QY	1694	GCTTCAACCGGCGGAAGATGAGGTCCCTACCGCTGAGGAGATGTCAGTTGACGTCAG		1753
Db	1522	GCTTCCACCGCGGAACTGCNACTTCCGAGCGCTGGAGGAGATGTCACACTGACACAG		1581
QY	1754	AGTCTCTCCCGAGCTGGCAAAACGCTCTGTTGCGGGAACCTTCACTCTTGGACAAAG		1813
Db	1582	AGTGTCTCCCGAGCTGGCGAAGAGTCTGTTGGGAACCTTCACTGAGCTGGAGAGG		1641
QY	1814	AAGAACAAATTTCTCTGCTGCTAAAGACAACTCTCAGCAGCATCAAGACACATCG		1873
Db	1642	AGGAGCAGATCTTGCTGTCATCAAAAGCAAACTCTGAGCTCCATCAAGCTGACATCG		1701
QY	1874	TCCATGCTTCTGTCGATGCCCGCGGAGTCAAGTGTGCTGACAGCCAGCTTCA		1933
Db	1702	TGCAGCTTCTCTGTCGATTTCCAGTCTCAGCCAGCGTCACTTCCAAACGAGCTTCC		1761
QY	1934	GGGCGGAGTACAAAGCCAGTGGCGGCCCTCCCTCTTCCAAAAGCCCGCTTCCAGG		1993
Db	1762	GGGCGGAGTACAAAGCCAGCGGGGGCGACGCGTTCAGAAACCGGCTCAAGTTCAGG		1821
QY	1994	TGACATCAGCTCTCTGAGGGGTCCAGACCCCTCCCGCGACGGGACGCGACGGAGGTG		2053
Db	1822	TGTGATATCACTTACAGCAGGCGT-----	GGGAGGCGCCACAAAGAGA	1863
QY	2054	GTCGCATTAATCCGTACACTTCACTCATCTGGGGTCCACGCGCGGTTCAAAGGAG		2113
Db	1864	ACGGCATCTACTCCGTCACTTCACTCCCTGCTCAGGCCCCAGCGCGGTTCAAAGGGG		1923
QY	2114	TGGTGGAGACCATTCAGGACAGCAGCTCCGAGACACTCAATGACACAGCCCTCCGTGAGGACC		2173
Db	1924	TGGTGGAGACCATTCAGGACAGCCAGCTGCTGAGACACACAGACCCGCTGCGGCCACGACT		1983
QY	2174	TGGCAGA	2180	
Db	1984	TGTCAGA	1990	
RESULT 5				
AD34315				
ID	AD34315	standard; cDNA; 2647 BP.		
XX	AD34315;			
XX	AC			
XX	16-JUL-2002	(first entry)		
DT	XX			
XX	XX			
DE	Human PKIN-18 cDNA.			
XX	Human; kinase; enzyme; PKIN-18 protein; immune system disorder; anaemia;			
KM	acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease;			
KM	asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease;			

KW AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma;
 KM leukemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder;
 KW Down's syndrome; gene therapy; protein therapy; cytosolic; gene; ss.
 XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 1.1995
 FT /tag- a
 FT /product- "Human PKIN-18 protein"

MO200218557-A2.

07-MAR-2002.

31-AUG-2001; 2001MO-US27219.

31-AUG-2000; 2000US-229873P.

08-SEP-2000; 2000US-231357P.

14-SEP-2000; 2000US-232654P.

22-SEP-2000; 2000US-234902P.

29-SEP-2000; 2000US-236499P.

06-OCT-2000; 2000US-238389P.

13-OCT-2000; 2000US-240542P.

(INCY-) INCYTE GENOMICS INC.

Bandman O, Nguyen DB, Walla NK, Hafalia AJA, Yao MG, Gandhi AR;
 Gururajan R, Ding L, Patterson C, Yue H, Baughn MR, Tribouley CM;
 Thornton M, Elliott VS, Lu Y, Ison CH, Au-Young J, Tang YT;
 P1 Azimtal Y, Burdill JD, Marcus GA, Zingler KA, Lu DM, Lal PG;
 P1 Rankmar J, Warren BA, Kearney L, Policky JL, Thangavelu K;
 P1 Burford N;

WPI: 2002-329769/36.
 P-PSDB: AAE21723.

New human kinases, useful for diagnosing, treating or preventing immune
 system disorders (e.g. Crohn's disease), neurological disorders (e.g.
 epilepsy), or cell proliferative disorders (e.g. cancers such as
 leukemia or lymphoma)

Claim 97, Page 212-213; 218pp: English.

The present invention relates to human kinases (PKIN) and polynucleotides
 encoding such proteins. PKIN sequences of the invention are useful for
 diagnosing, treating or preventing disorders associated with aberrant
 expression of PKIN, particularly immune system disorders (e.g. acquired
 immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,
 anaemia, asthma), neurological disorders (e.g. epilepsy, Charcot-Marie-
 Tooth disease or seizures), cell proliferative disorders (e.g. cancers
 such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma),
 and developmental disorders (e.g. Down's syndrome). They are also used
 in gene therapy and protein therapy. The present sequence is a cDNA
 encoding human PKIN-18 protein.

Sequence 2647 BP; 525 A; 885 C; 789 G; 448 T; 0 other;

Query Match 41.0%; Score 955.8; DB 24; Length 2647;

Best Local Similarity 70.0%; Pred. No. 1.6e-197;

Matches 1473; Conservative 0; Mismatches 437; Indels 195; Gaps 6;

QY 88 TATGTGGGCCCCATATCGCTGGAGAGACGCTGGGCAAGAGACAGACAGGCTGTAA 147
 DB 31 TATGTGGGCCCCATATCGCTGGAGAGACGCTGGGCAAGAGAGACAGGCTGTAA 90
 QY 148 CTGGGGGTCACGATCAGAGGCTCAGAGGTCGCGATCAGATCTGAACCGGAGAG 207
 DB 91 CTGGGGGTCACGATCAGAGGCTCAGAGGTCGCGATCAGATCTGAACCGGAGAG 150
 QY 208 CTGCTGAGTCTGCTGATGAGAGGAGAGGAGAGATCGCATCTGAGCTATCGAA 267
 DB 151 CTGAGGAGTCTGCTGATGAGAGGAGAGGAGAGATCGCATCTGAGCTATCGAG 210

QY 268 CACCACATGCTCTCAAGCTCCAGAGCTCTACGAGAACAGAAATATTGTACCTGTT 327
 DB 211 CACCCACAGCTCTCAAGCTCTACGAGAGCTTATGAAACAAATAATTTGTACCTGTT 270
 QY 328 CTGAGAGAGCTCTGGGGGGGAGAGCTATTCGATCTCTGATTAAGAGAGAGAGTGA 387
 DB 271 CTGAGAGAGCTCTGGGGGGGAGAGCTTCTGATCTCTGATTAAGAGAGAGAGTGA 330
 QY 388 CCGAGAGAGGCCCCGAAAGTTCTTCCGAGATGATGCTGGCTGAGCTTCCGACAGC 447
 DB 331 CCGAGAGAGGCCCCGAAAGTTCTTCCGAGATGATGCTGGCTGAGCTTCCGACAGC 390
 QY 448 TACTCATCTGCGACAGAGACTTAAGCCGAGAACTGCTTTTGGATGAGAAAAAAC 507
 DB 391 CACTCATATGCGACAGAGAGATCTGAACCTGAAACCTCTGCTGAGAGAGAAACAC 450
 QY 508 ATCCGATTCAGAGCTTCGCGATGCGCTCCCTGAGAGTGGGGAGACGCTCTGAGAGC 567
 DB 451 ATCCGATTCAGAGCTTCGCGATGCGCTCCCTGAGAGTGGGGAGACGCTCTGAGAGC 510
 QY 568 AGCTGCGGTCGCCCATTTATGCTGTCAGAGAGGATTAAGGGGAGAAATATGATGAC 627
 DB 511 AGCTGCGGTCGCCCATTTATGCTGTCAGAGAGGATTAAGGGGAGAAATATGATGAC 570
 QY 628 CGCCGGGACAGATGCTGAGAGCTGTGAGATCTCTTTCGCTGCTGAGGGGCTCTG 687
 DB 571 CGGAGAGGCGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTG 630
 QY 688 CCTTTGATACGACAACTCCGAGCTGTGAGAGAGTGAAGAGGGGCTCTTCCAC 747
 DB 631 CCTTTGATACGACAACTCCGAGCTGTGAGAGAGTGAAGAGGGGCTCTTCCAC 690
 QY 748 ATGCCCATCTTCAATCTCCAGATTCGAGAGCTCTTCCGAGAGAGTATGAGTGA 807
 DB 691 ATGCCCATCTTCAATCTCCGAGATTCGAGAGCTCTTCCGAGAGAGTATGAGTGA 750
 QY 808 CCGGAAAAAGGCTCACTGTGAGCAAAATTCAGAAATCTTGTATCAGGGGGAA 867
 DB 751 GCGGACGCGCGCTCAGCTGAGACATTCAGAAATCATATGATATAGGGGGCAAG 810
 QY 868 CAGAGCGCAAGCCGCTCTGAGAGCCAGCCCTGCGGGGAGAGCATGCGAGAGCTG 927
 DB 811 AATGAGCCGCAAGC-----AGAGCAGCCATCTCTGCAAGAGTGCAGATCGCTG 864
 QY 928 CCATCCAGAGAGAGCTGAGAGCCGAGCTCTAGAGAGAGAGATCAGTGGCTGTT 987
 DB 865 CCAGCTGAGAGAGAGCTGAGAGCCGAGCTCTAGAGAGAGAGATCAGTGGCTGTT 924
 QY 988 AGGAGCCGAGAGAGCTGAGAGCTGCGAGCTGCGAGAGAGAGAGAGAGAGATG 1047
 DB 925 CGAGACCGCAAGAGCTGCTGAGAGAGCTGCTGCGAGAGAGAGAGAGAGAGATG 984
 QY 1048 ATATATTTATGCTTTTGGATTCGAGAGAGAGAGAGAGAGATTCAGAGTGAAGAGAGAGCTG 1107
 DB 985 ATTTACTTCTCTCTCTGAGAGAGAGAGAGAGAGAGATTCAGAGTGAAGAGAGAGCTG 1044
 QY 1108 CCTCCCGAGATGATGTTGAGAGCCCGGAGAGAGAGAGATTCAGATTCAGAGAGCTG 1167
 DB 1045 CCGCCCGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1104
 QY 1168 CACGGAAG 1227
 DB 1105 CACGGAAG 1155
 QY 1228 GGTGTGTGCTCCCTGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1287
 DB 1156 GACGAG 1215
 QY 1288 TCCGATGAG 1347
 DB 1216 TCGGAGTCAATGAG 1270

QY	1348	AGTCGGGCTTTTCCCTTTTTCACCGGAGCGGGGGGTGAGATGAGGCTCGAGCGGGGGGC	1407
Db	1271	-----	1270
QY	1408	TCCCCGACTTCCAAAACGACGACGCTGCTTTCGgggGCCCAAGGGTGGGGGCGCGGG	1467
Db	1271	-----	-CCGGGT 1277
QY	1468	GAGCAGCCCCCGCCCCAGTGCCCGCTCACACCCTGCCCCGGCCCCCAAGCTTCCC	1527
Db	1278	GACCCCTACCCCTTACCAAGAGGGGCGTCCCTCCCTC-----	1314
QY	1528	CGCTCTGGGGGGGACCCCGCTTGCACTGGCGCTGACACAGCCCGGGCCAGTCCACC	1587
Db	1315	-----ACCCCAAAGGGGACAACCTGTCCACAGGCCAAAGAGAGCGCGCT	1359
QY	1588	GGGACCCCGGGGACAACACACCCCGCCAGCCCGGGGGTGGCGTGGGGGAGCGCGTGG	1647
Db	1360	GGCAGGCCAACCCACAGCGCCCCCGTCCAGGCC-----AGCGTGGAGGGGTCCCTCG	1413
QY	1648	AGGAGTCTCTCACTCATTCGCGAAGCTTCTGGGCTCCCTCGCTTTCACCGCGCG	1707
Db	1414	AGGGGCGCGCTCACTCCATCAAGAAGAGCTTCTGGGCTCACCCCTTCCACCGCGGG	1473
QY	1708	AAGATGAGGTCCCAACGCGCTGAGGAGATGTCCAGCTTGAGCGCAGAGTCTTCCCCGGAG	1767
Db	1474	AAATCTCAAGTTCGAGCGCCGGAGGAGATGTCCAACCTGACACCAAGATGTCTCCCAAG	1533
QY	1768	CTGGCAAAAGCGCTCTGTGTTCGGGAATTCTCTTGTGACAAAGAACAATAATTC	1827
Db	1534	CTGGGGAAGAGTCTGTGTGGGAATTCTCTTGTGAGAGTCTGAGAGAGAGCAATCTTC	1593
OY	1828	CTCTGCTTAAGAGCAAACTCTCAGCAGCATCAAAAGACAGACATGTCATGCTTCTG	1887
Db	1594	GTGTCTATCAAAAGACAACCTCTGAGCTCATCAAGAGCTGACATCGACAGCTTCTG	1653
QY	1888	TGCATCCCCAGCGCTAGTACAGTGTGTGTACAGACCAAGCTTCAGGGCCGAGTACAAG	1947
Db	1654	TGATTTCCAGTCTAGCCACAGCGTCATCTCCAAACGAGCTTCGCGGCGAGTACAAG	1713
QY	1948	GCCAGTGGCGGCGCTCTCTTCCAAAAGCCCGCTCGCTCCAGTGTGAGCATGAGCTCC	2007
Db	1714	GCCACGGGGGGGCGACCGCTGTTCAGAAACCGGTCAAGTTCACAGTTGATTAACCTAC	1773
QY	2008	TCTGAGGCTCCAGAGCCCTCCCGCGAGCGGAGCGGACGCGAGGTTGGTGGCATCTACTC	2067
Db	1774	ACGGAGGGT-----GGGAGGCGGAGGAAGGAACGGGATCTACTCC	1815
QY	2068	GTCACCTTCACTCTCATCTCGGGGTCCAGCGCGTGGTTCAAAGGATGTGTGAGACATC	2127
Db	1816	GTCACCTTCACTCTCTCTCTCAGGCCCCCAGCCCGTCTTCAAAGGGGTGTGTGAGACCTATC	1875
QY	2128	CAGGACAGACTCTTAGAGCACTATGACACAGCCCTCGGTGAGAGGCCCTGGGACGAGAGAG	2187
Db	1876	CAGGCGCAGAGCTGTAGCACACAGACCCGCGCTGCGGCCACGACCTTGTACAGACCACT	1935
QY	2188	AACGG 2192	
Db	1936	AAC TG 1940	
RESULT 6			
XX	ABA08296		
ID	ABA08296 standard; cDNA; 906 BP.		
XX	ABA08296;		
AC			
DT	11-JAN-2002 (first entry)		
DE			
XX	Human HRPOLK-1 homologue-encoding cDNA, SEQ ID NO:72.		
KW	Human; cytokine; cell proliferation; cell differentiation; growth factor		
KM	haematopoiesis regulation; tissue growth; immunomodulator; activin;		

KW inhibit chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytosolic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antilucer; ss.
 XX
 OS Homo sapiens.
 PN WO200157188-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US03800.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PT Tang YT, Liu C, Dirmannac RT;
 XX
 DR WPI: 2001-457740/49.
 DR P-PSDB; ABB11052.
 XX
 PT Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 XX
 PS Claim 1; Page 364-365; 1963pp; English.
 CC Sequences ABB10961-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention.

XX Sequence 906 BP; 205 A; 253 C; 267 G; 181 T; 0 other;
 SQ Best Local Similarity 28.2%; Score 658.8; DB 22; Length 906;
 Query Match Pred. No. 2.9e-133;
 Matches 786; Conservative 0; Mismatches 2; Indels 106; Gaps 2;

```

79 CACGCCCAATATGTGGGCGCCCTATCGGCTGAGAAAGCGCTGGGCAAGAGACAGACAGG 138
DB 1 CCGCCCAATATGTGGGCGCCCTATCGGCTGAGAAAGCGCTGGGCAAGAGACAGG 60
QY 139 CCGGTTAAATGTGGGCGCCCTATCGGCTGAGAAAGCGCTGGGCAAGAGATCGTGAAC 198
DB 61 CCGGTTAAATGTGGGCGCCCTATCGGCTGAGAAAGCGCTGGGCAAGAGATCGTGAAC 120
QY 199 CCGGAGAAAGCTGTGAGAGTGTGCTGATGAGAGTGGAGGAGATCGCATCT-GAA 257
DB 121 CCGGAGAAAGCTGTGAGAGTGTGCTGATGAGAGTGGAGGAGATCGCATCTCGAG 180
QY 258 GCTCATGGAACACCCACATGTCTCAAGCTCCACGACGCTTACGAGAAAGAAATATTT 317
DB 181 GCTCATGGAACACCCACATGTCTCAAGCTCCACGACGCTTACGAGAAAGAAATATTT 240
QY 318 ----- 317
DB 241 TCCCGCGGATGAACATGACATCAGGTCGCTGAGTGGCGGAGGTTTCACACAGGGAA 300
QY 318 -----GTACTGTGTTCTGGA 332
DB 301 GCTGTGTGCGCGCGGCTCATGAGACCTGCTCAGGCTTCCAGATACCTGTCTTGA 360
QY 333 GCACGCTCTCGGGGGGTGAGTATTCATGACTACCTGTTAAAGAGGGAGACGTACGCCAA 392
DB 361 GCACGCTCTCGGGGGGTGAGTATTCATGACTACCTGTTAAAGAGGGAGACGTACGCCAA 420
QY 393 GCGAGGCGCGAAATGCTTCCGCAAGTGTGCTGCGCTGACACTTTCGCCACAGTACTC 452
DB 421 GCGAGGCGCGAAATGCTTCCGCAAGTGTGCTGCGCTGACACTTTCGCCACAGTACTC 480
QY 453 CATCTGCCACAGAGACTTAAAGCCCGAGAACTGTTTGGATGAGAAACAAACATCCG 512
DB 481 CATCTGCCACAGAGACTTAAAGCCCGAGAACTGTTTGGATGAGAAACAAACATCCG 540
QY 513 CATCTGCCACAGACTTGGCATGCGCTCCCTGCAAGGTGGGAGACACCTCTGAGACCACTG 572
DB 541 CATCTGCCACAGACTTGGCATGCGCTCCCTGCAAGGTGGGAGACACCTCTGAGACCACTG 600
QY 573 CCGGTCGCCCATTTATGCGGTCCAGAGGTGATTAAAGGGGAGAAATATGATGGCGCG 632
DB 601 CCGGTCGCCCATTTATGCGGTCCAGAGGTGATTAAAGGGGAGAAATATGATGGCGCG 660
QY 633 GCGAGACATGTGAGAGTGTGAGTCACTTCCGCTGCTGAGGGGCTGCGCTT 692
DB 661 GCGAGACATGTGAGAGTGTGAGTCACTTCCGCTGCTGAGGGGCTGCGCTT 720
QY 693 TGTATGACGACACCTCCGCCAGCTGCTGAGAAAGGTGAAAGGGGCGCTTCCACATCC 752
DB 721 TGTATGACGACACCTCCGCCAGCTGCTGAGAAAGGTGAAAGGGGCGCTTCCACATCC 780
QY 753 CCACTTCATTCCTCCAGATTCGACAGCCCTCCTGAGGGGATGATGAGAGGACCCGA 812
DB 781 CCACTTCATTCCTCCAGATTCGACAGCCCTCCTGAGGGGATGATGAGAGGACCCGA 840
QY 813 AAAAAGCTCAGTCTGAGAGCAATTCAGAAACATCTTGGTACTAGCGGGAA 866
DB 841 AAAAAGCTCAGTCTGAGAGCAATTCAGAAACATCTTGGTACTAGCGGGAA 894

```

RESULT 7
 ABA08924
 ID ABA08924 standard; cDNA; 614 BP.
 AC ABA08924;

XX 11-JAN-2002 (first entry)
 DT Human Ser/Thr kinase homologue-encoding cDNA, SEQ ID NO:700.
 DE Human; cytokine; cell proliferation; cell differentiation; growth factor;
 XX haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokines; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antihistaminic; antihistric; haemostatic; antiarteriosclerotic;
 KW cytosstatic; osteopathic; vasotropic; cardiac; virucide; antibacterial;
 KW antifungal; vulnery; antitumor; ss.

XX Homo sapiens.
 OS WO200157188-A2.
 PN 09-AUG-2001.
 PD 05-FEB-2001; 2001WO-US03800.
 PF 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 XX (HSE-) HSEQ INC.
 PA Tang YT, Liu C, Drmanac RT;
 PI MPI: 2001-457740/49.
 XX P-PSDB; ABB11680.
 DR Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 XX Claim 1; Page 648; 1963pp; English.

XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention. Methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides,
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to

CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention.

XX Sequence 614 BP; 128 A; 226 C; 161 G; 99 T; 0 other;

Query Match 26.3%; Score 614; DB 22; Length 614;
 Best Local Similarity 100.0%; Pred. No. 1.3e-123;
 Matches 614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1718 TCCCTACCCGCTGAGAGATGCTCCAGCTTGACGCGCAGAGTCTCCCGAGCTGGCAAAAC 1777
 DB 1 TCCCTACCCGCTGAGAGATGCTCCAGCTTGACGCGCAGAGTCTCCCGAGCTGGCAAAAC 60
 QY 1778 GCTCTGAGTGGGAACTTATCTCTGAGCAAGAAATATCTCTGAGTAA 1837
 DB 61 GCTCTGAGTGGGAACTTATCTCTGAGCAAGAAATATCTCTGAGTAA 120
 QY 1838 AGGACAACTCTCAGAGCATCAAGACATGCTCATGCTTCTGTCATCCCA 1897
 DB 121 AGGACAACTCTCAGAGCATCAAGACATGCTCATGCTTCTGTCATCCCA 180
 QY 1898 GCTGAGTACAGTGTGCTGTACACAGCAAGCTTCAAGGCGGAGTACAGGCGCACTGCG 1957
 DB 181 GCTGAGTACAGTGTGCTGTACACAGCAAGCTTCAAGGCGGAGTACAGGCGCACTGCG 240
 QY 1958 GCGCCCTCCGCTTCAAAAGCCGCTGCTCCAGGTGAGCATAGCTCTCTGAGGCTC 2017
 DB 241 GCGCCCTCCGCTTCAAAAGCCGCTGCTCCAGGTGAGCATAGCTCTCTGAGGCTC 300
 QY 2018 CAGAGCCCTCCGCGAGGAGGAGGAGGAGGAGGAGTGTGATCTACCTGACCTTCA 2077
 DB 301 CAGAGCCCTCCGCGAGGAGGAGGAGGAGGAGGAGTGTGATCTACCTGACCTTCA 360
 QY 2078 CTCTCATCTGCGGTCCAGCGCTGCTTCAAGCAGTGTGTGAGAACATCCAGGCGACAGC 2137
 DB 361 CTCTCATCTGCGGTCCAGCGCTGCTTCAAGCAGTGTGTGAGAACATCCAGGCGACAGC 420
 QY 2138 TCTGTAGCACTATGAGCAAGCCCTGCTGCAAGGCGCTGAGAGAGAAAGGCGGCC 2197
 DB 421 TCTGTAGCACTATGAGCAAGCCCTGCTGCAAGGCGCTGAGAGAGAAAGGCGGCC 480
 QY 2198 AGACCCGCGCTGCTGTGTCGCCACCCCGAAAGCTGTGACGCCCGCGCGCAGAC 2257
 DB 481 AGACCCGCGCTGCTGTGTCGCCACCCCGAAAGCTGTGACGCCCGCGCGCAGAC 540
 QY 2258 CAGAGCTGAGCACTGTCCCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2317
 DB 541 CAGAGCTGAGCACTGTCCCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
 QY 2318 ACGGAGCCCTCTG 2331
 DB 601 ACGGAGCCCTCTG 614

RESULT 8
 ABK70216
 ID ABK70216 standard; cDNA; 512 BP.

XX ABK70216;

XX 15-JUL-2002 (first entry)

XX Human lung cancer associated cDNA SEQ ID 87.

XX Human; ss; gene; lung cancer; cytostatic; tumour; vaccine.

XX Homo sapiens.

XX

PN WO200224057-A2.
 XX 28-MAR-2002.
 PD 20-SEP-2001; 2001WO-US42232.
 PP 22-SEP-2000; 2000US-234837P.
 PR 10-OCT-2000; 2000US-239440P.
 PR 29-JUN-2001; 2001US-301928P.
 XX (CORI-) CORIXA CORP.
 PA Benson DR, Mohamath R, Lodes MJ;
 PI WPI; 2002-372001/40.
 PS New tumour lung proteins and nucleic acids encoding the proteins, useful
 PT as vaccines and for treating, preventing, diagnosing or monitoring lung
 PT cancer.
 XX Claim 1; Page 132-133; 189pp; English.

The invention relates to an isolated polynucleotide comprising a sequence
 selected from 183 human DNA sequences (appearing as ABK70130-ABK70312),
 or their fragments, homologues, variants or complements and their encoded
 polypeptides. Also included are an expression vector comprising the
 polynucleotide operably linked to an expression control sequence; a host
 cell transformed or transfected with an expression vector of; an isolated
 antibody, or its antigen-binding fragment that specifically binds to the
 polypeptide; a method for detecting the presence of a cancer in a
 patient; a fusion protein comprising at least the polypeptide; an
 oligonucleotide that hybridises to the polynucleotide under moderately
 stringent conditions; a method for stimulating and/or expanding T cells
 specific for a tumour protein; an isolated T cell population comprising T
 cells prepared from the method of above; a composition comprising a first
 component consisting of carriers and immunostimulants, and a second
 component selected from the polynucleotides, proteins, antibodies, fusion
 proteins, T cell populations and antigen presenting cells expressing the
 polypeptide; methods for stimulating an immune response or treating
 cancer in a patient by administering the composition and diagnostic kits
 comprising at least one of the oligonucleotide of, or an antibody and a
 detection reagent consisting of a reporter group. The polypeptides and
 polynucleotides are useful as vaccines and for the treatment or prevention of
 lung cancer, and for diagnosis and monitoring of such cancer. The
 polynucleotide, polypeptide and antigen presenting cells can be
 used to stimulate or expand T cells specific for a tumorous protein.
 CC The polynucleotides may be used as probes or primers for nucleic acid
 CC hybridisation, and in the preparation of ribozyme molecules for
 CC inhibiting expression of tumour polypeptides and proteins in tumour
 CC cells. The present sequence is one of the 183 lung cancer associated
 CC polynucleotides.

XX Sequence 512 BP; 108 A; 149 C; 166 G; 89 T; 0 other;

Query Match 21.9%; Score 512; DB 24; Length 512;
 Best Local Similarity 100.0%; Pred. No. 1.6e-101;
 Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 519 AGACTTCGCGATGCGGCTCCGTCAGGTGGGGGACAGCTCTGTGAGACAGCTGGGCTC 578
 DB 1 AGACTTCGCGATGCGGCTCCGTCAGGTGGGGGACAGCTCTGTGAGACAGCTGGGCTC 60
 QY 579 CCCCATTATGCTGTCCAGAGGTGATTAAGGGGAAATATGATGAGCCCGCGGCGA 638
 DB 61 CCCCATTATGCTGTCCAGAGGTGATTAAGGGGAAATATGATGAGCCCGCGGCGA 120
 QY 639 CATGTGAGCTGTGAGTATCTCTTCCGCTGCTGTGGGGGCTGCGCTTTGATGA 698
 DB 121 CATGTGAGCTGTGAGTATCTCTTCCGCTGCTGTGGGGGCTGCGCTTTGATGA 180
 QY 699 CGACAACCTCCGCGAGCTGTGAGAGTGAAGAGGGGGGCTCTTCACATGAGCCCACTT 758
 DB 181 CGACAACCTCCGCGAGCTGTGAGAGTGAAGAGGGGGGCTCTTCACATGAGCCCACTT 240

QY 759 CATCTCTCAGATTGCGAGACCTCTGAGGGAATGATCAAGTGAGCCCGAAAAAG 818
 DB 241 CATCTCTCAGATTGCGAGACCTCTGAGGGAATGATCAAGTGAGCCCGAAAAAG 300
 QY 819 GCTCAGTCTGAGCAATATGAGAAACATCTCTGATCTAGCCGGGAAACAGACCCAGA 878
 DB 301 GCTCAGTCTGAGCAATATGAGAAACATCTCTGATCTAGCCGGGAAACAGACCCAGA 360
 QY 879 CCCGTCCTGAGACCCAGCCCTGCGCCCGGATAGCCATGCGGAGCCCTGCATCCAGG 938
 DB 361 CCCGTCCTGAGACCCAGCCCTGCGCCCGGATAGCCATGCGGAGCCCTGCATCCAGG 420
 QY 939 AGAGCTGAGACCCGAGCTCTAGAGACATGAGCATCACTGGCTCTCAGGAGCCGCGA 998
 DB 421 AGAGCTGAGACCCGAGCTCTAGAGACATGAGCATCACTGGCTCTCAGGAGCCGCGA 480
 QY 999 GAGGCTGATCGGAGCTGCGGCACTGAGGAG 1030
 DB 481 GAGGCTGATCGGAGCTGCGGCACTGAGGAG 512
 RESULT 9
 ABL10489
 ID ABL10489 standard; cDNA; 2720 BP.
 AC ABL10489;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 25949.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmacological; gene; ss.
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 DR WPI; 2001-656860/75.
 DR P-PSDB; ABB6386.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Claim 1; SEQ ID NO 25949; 21bp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB16176-AB16176), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB557737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 2720 BP; 661 A; 715 C; 759 G; 585 T; 0 other;

Query Match 20.2%; Score 470.8; DB 23; Length 2720;
 Best Local Similarity 66.4%; Pred. No. 2e-92;
 Matches 692; Conservative 0; Mismatches 347; Indels 3; Gaps 1;
 QY 77 AGCAGCCCAATATGTCGGCCCTATCGGCTGAGAGACGCTGGGAAAGAGACAG 136
 DB 29 AGATTGCCAATTTGGGGCCCTATCGCTGAGAAACCTTGGGCAAGGGTAAACCG 88
 QY 137 GCGTGTAAACTCGGGGTCACATGATCAGGGGTGAAGAGTGCATCAAGATCGTGA 196
 DB 89 GTCTGTCAGAGTTGGGCGTCAATGTGTATGTGGCAAGAGGTGGCATTAATAATCA 148
 QY 197 ACCGAGAGACCTGTGCGAGTGGTGTGATGAGGTGAGCGGAGATCGCATCTGA 256
 DB 149 ATCGCGAGAAACTCAGCGAATCGGTGCTAATGAAGGTGAACGTAATGCCATTAAT 208
 QY 257 AGCTCATGACACCCACATGCTCTCAAGTCCACGACGCTACAGAAACAAAGAAATTT 316
 DB 209 AACTAATCGATCAATCCACACGCTCTTGGCTGAGCCGATGTGTACGAAACAAAGATATT 268
 QY 317 TGTACTGTTCTGAGACACGCTCGGGGGGTAGATTAATGACTACTGTTAAAGAAAG 376
 DB 269 TGTATTGATATTGAGACATGATCCGGCGGAGAGCTTCTGATTAAGTGAAGAG 328
 QY 377 GGAGACTGACGCCCAAGAGAGCGCGAAAGTTCTCCGACAGATTGTCTGCGCTGACT 436
 DB 329 GTGCATTTGAGCGCGAAGAGAGCGCGCAAGTTCTCAGGCAAAATCATCTCCGCCCTGATTT 388
 QY 437 TCTGCGACAGCTACTCCATCTGACAGACAGACCTAAAGCCCGAAGACCTGTTTGGATG 496
 DB 389 TCTGCCACTGCGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 448
 QY 497 AGAAAAACAACATCCGATTCGAGACTTCCGATGCGCTCCGCTGAGTGGGAGACACC 556
 DB 449 AGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 508
 QY 557 TCTGAGACACAGCTCGGGTCCCGCATTAATGCTGTCCAGAGGTATTAAGGGGAAA 616
 DB 509 TGTGGAGACCTCTCGGGAGCCCACTACGCTGTCCAGAGGTATTAAGGGGCGAGA 568
 QY 617 AATATGATGGCGCGGGGAGACATGAGAGTGTGAGATCTGTCCTGCTGCTGCTGCTG 676
 DB 569 AGTACATGGCGCGAAGCGGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 628
 QY 677 TGGGGCTGCTGCTTGTATGAGACAACTCCGACGCTGTGAGAAAGTGAAGGG 736
 DB 629 TGGGTGCTTGGCTTGAAGAGCAAACTTGGGCGAGCTGTGAGAAAGTGAAGGG 688
 QY 737 GCGTCTTCAATGCCCCCACTTCAATCTCCAGATTCGACAGAGCTCTGAGGGAAATGA 796
 DB 689 GCGTCTTCAATGCCCCCACTTGTGCGCGGAGCTCCAGAGCTCTGCGCGGATGA 748
 QY 797 TCGAAGTGGAGCCGAAAAAAGGCTCAGTGTGAGCAAAATTCAGAAACATCTTGGTAA 856
 DB 749 TTGAGTCAATTCGAGACCGCGGCTCAGCTGTGTAATCAACCGGATCTGCGGTCA 808
 QY 857 TAGCGGAGAAACAGACGACCCGCTGAGAGCCAGCCCTGCGCGGCTGATACCA 916
 DB 809 CAGCTGGCGGAAAGG---GAGCTGAGGTGAGTGCAGCAATGATGAGAGTGTGAGA 865
 QY 917 TCGGAGCCTGCCATCCACAGAGAGCTGAGACCCGACGCTTGAAGAGATGCAATCAC 976
 DB 866 CACACCTTATTCACAGCAGCAGCGTGTGATCCGATGTGTGAACGCGATTTGCTGCG 925
 QY 977 TGGGCTGTTCAAGGAGCGGAGAGGTGATGCGAGCTGCGAGTGAAGAGAGAGAAC 1036
 DB 926 TGGGCTGTTCAAGGAGAGAGAGAACTCAACAGGAAGCTCAGTCAAGTCAATGACATA 985
 QY 1037 AAGAAAGATATATATATATCTCTTGTGATGGAAGAGAGCGTATCCAGCTGATAG 1096
 DB 986 CGGAGAGATTAATATTTCTGCTGCTGAGGCAAAAGACAGACCTGCGCTGAGAG 1045
 QY 1097 ACCAGAGCTGCTCCCGGAA 1118

PR 17-MAY-2000; 2000DE-1024171.
 (BADI) BASF-LYNX BIOSCIENCE AG.
 PI Schneider A, Klausner B, Fischer A, Newrzella D, Goetz B;
 PI Rosner M, Eisenhardt G, Kuner R, Trutzel A, Kammandel B;
 PI Jomana Naim S, Schwaninger M;
 XX WPI: 2002-055696/07.
 DR P-PSDB; ABB04433.
 XX
 PT New polynucleotide, useful for the diagnosis, treatment and prevention
 PT of e.g. apoplexy, tumors and autoimmune disease, comprises a
 PT polynucleotide encoding the neuronal protein kinase -
 PS
 PS Claim 1; Page 54-55; 75pp; German.
 CC The present invention provides the protein and coding sequences of two
 CC neuronal serine threonine protein kinases each from the mouse and human.
 CC These are designated 9B5 and 9B5_b. Detection of their expression levels
 CC can be used to diagnose the risk of apoplexy and in prognosis of cerebral
 CC infarction. They can also be used to identify agents useful in the
 CC prevention, treatment and diagnosis of neurological diseases, including
 CC stroke, multiple sclerosis, Parkinson's disease, amyotrophic lateral
 CC sclerosis, tumours, especially carcinoma, immunological disorders,
 CC including autoimmune diseases, atopy, viral (including human immune
 CC deficiency virus) infections, acute or chronic leukocytic or myeloid
 CC leukaemia, primary chronic polyarthritis, Crohn's disease and ulcerative
 CC colitis. The present sequence is the human 9B5 coding sequence.
 CC
 XX
 XX Sequence 3312 BP; 710 A; 1071 C; 954 G; 577 T; 0 other;
 SO
 Query Match 11.3%; Score 264.2; DB 24; Length 3312;
 Best Local Similarity 57.6%; Pired. No. 1.3e-47;
 Matches 473; Conservative 0; Mismatches 348; Indels 0; Gaps 0;
 32 GCTCTCCCGCTACACACTGCCCCACCCACCCACCCACCCACCCACCCCAATATG 91
 170 GCTCTAGGCTGCGCGTGGCGGAACTCCTCTGTCGAGAGAGAGAGAGAGAGAG 229
 92 TGGGCCCCCTATCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 151
 230 TGGGCAACTGACCCCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 289
 152 GGGTCACATGACAG 211
 290 CTGGGACATCTACAG 349
 212 CGGAGTCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 271
 350 ATCCAGACAGCTGAG 409
 272 CACATGCTCTGAG 331
 410 CCAACATCTGAG 469
 332 AGCAGCTCTGAG 391
 470 AGTACCAAGTGTGAG 529
 392 AG 451
 530 AGGAG 589
 452 CCAATGCTGAG 511
 590 ATATGCTGAG 649
 512 GCATGAG 571
 650 AGATGCTGAG 709
 572 GCGGAG 631

DB 710 GCGGAG 769
 632 GGGCAGACATGAG 691
 770 AGGTGACATCTGAG 829
 692 TTGATGACAG 751
 830 TCGAGCGGAG 889
 752 CCCATCTGATCTGAG 811
 890 CTCTTACATGTAACAG 949
 812 AAAAAAGCTGAG 852
 950 CTAAGCGTGTACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 990
 DB
 RESULT 13
 ABA05740
 ID ABA05740 standard; cDNA; 3392 BP.
 XX
 AC ABA05740;
 XX
 DT 04-MAR-2002 (first entry)
 XX
 DE Human neuronal serine threonine protein kinase 9B5_b coding sequence.
 XX
 KW Human, neuronal serine threonine protein kinase; apoplexy; cyrostatic;
 KW cerebroprotective; neuroprotective; antiparkinsonian; immunosuppressive;
 KW virucide; anti-HIV; antitubercular; antitumor; cancer;
 KW cerebral infarction; neurological disease; immunological disease; atopy;
 KW infection; leukaemia; polyarthritis; Crohn's disease; ulcerative colitis;
 KW 9B5_b; ss.
 XX
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 64..2130
 FT /tag= a
 FT /product= "9B5"
 FT 335..3360
 FT /tag= b
 XX
 PN MO200188108-A1.
 XX
 PD 22-NOV-2001.
 XX
 PF 17-MAY-2001; 2001WO-BE05660.
 XX
 PR 17-MAY-2000; 2000DE-1024171.
 XX
 PA (BADI) BASF-LYNX BIOSCIENCE AG.
 XX
 PI Schneider A, Klausner B, Fischer A, Newrzella D, Goetz B;
 PI Rosner M, Eisenhardt G, Kuner R, Trutzel A, Kammandel B;
 PI Jomana Naim S, Schwaninger M;
 XX WPI: 2002-055696/07.
 DR P-PSDB; ABB04434.
 XX
 PT New polynucleotide, useful for the diagnosis, treatment and prevention
 PT of e.g. apoplexy, tumors and autoimmune disease, comprises a
 PT polynucleotide encoding the neuronal protein kinase -
 PS
 PS Claim 1; Page 55-56; 75pp; German.
 CC The present invention provides the protein and coding sequences of two
 CC neuronal serine threonine protein kinases each from the mouse and human.
 CC These are designated 9B5 and 9B5_b. Detection of their expression levels
 CC can be used to diagnose the risk of apoplexy and in prognosis of cerebral

CC Infarction. They can also be used to identify agents useful in the
CC prevention, treatment and diagnosis of neurological diseases, including
CC stroke, multiple sclerosis, Parkinson's disease, amyotrophic lateral
CC sclerosis, tumors, especially carcinoma, immunological disorders,
CC including autoimmune diseases, atopy, viral (including human immune
CC deficiency virus) infections, acute or chronic leukocytic or myeloid
CC leukemia, primary chronic polyarthritis, Crohn's disease and ulcerative
CC colitis. The present sequence is the human 985_b coding sequence.

XX Sequence 3392 BP; 726 A; 1099 C; 972 G; 595 T; 0 other;

Query Match 11.3%; Score 264.2; DB 24; Length 3392;

Best Local Similarity 57.6%; Pred. No. 1.3e-47;

Matches 473; Conservative 0; Mismatches 348; Indels 0; Gaps 0;

```

32 GCTCTCCCGCCGACCACTCCACCCACCCACCCACCCACCCACCCACCCATATG 91
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
170 GCTCACTGGGTGGCCGTCGCGAATCCATCGCTCTGTCGAGAGACAGCCACAG 229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
92 TGGGCCCTATGGGCTGGAGAGAGCTGGGCAAGAGACAGAGAGGCTGTTAACTCG 151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
230 TGGGCACTACCCGCTGCTAGAGACATTTGGGAAGGCAATTTGCCAAGTCAAGCTG 289
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
152 GGGTCACTGACATCAGAGGTCAGAGAGTGGCATCAGATCGTAACCGGAGAACTGT 211
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
290 CTCGGACATCTCACTGCTGGAGAGGTGCGCATCAAGATTATCGCAAAACCCAGCTGA 349
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
212 CGGAGTCGCTGATGATGAGTGGAGCGGAGATCGCATCTCAAGCTCATCAACAC 271
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
350 ATCCAGCAGACCTGAGAAAGCTGTTCGAGAGATCGCATGAGAGGCTTAACACAC 409
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
272 CACATGTCCCAAGCTCCAGAGAGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAG 331
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
410 CCACATCTCGGAAGCTCTTGTAGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 469
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
332 AGCAGCTCCGCGGGGAGAGCTATTGACTACCTGTAAGAGAGAGAGAGAGAGAGAG 391
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
470 AGTACCAAGTGTGAGAGAGTGTGTTGACTACCTGTTGCTGCTGCTGCTGCTGCTG 529
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
392 AGGAGCCCGGAAAGTCTTCCGCGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 451
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
530 AGGAGCTCGAGCGAGAGTTCGAGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 589
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
452 CCATCTGCCAGAGAGAGCTTAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 511
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
590 ATATTGTACACAGGAGAGCTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 649
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
512 GCATTCAGACTCGGAGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 571
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
650 AGATTGCTGACTTTGGCTTACAGAGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 709
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
572 GCGGTCCTCCCATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 631
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
710 GGGGGGCCCCCATATGCGCGCCGAGAGCTGTTTACAGGCAAGAGAGAGAGAGAGAG 769
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
632 GGGGAGACATGAGAGCTGAGAGTCACTCTTCCGCTGCTGCTGCTGCTGCTGCTGCT 691
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
770 AGGTGACATCTGAGAGCTGGAGATCTGTAACACCTGCTGAGGCGCTGCTGCTGCT 829
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
692 TTGATGACAGACACTCCGACAGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 751
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
830 TGGAGGGGCAACATCTCAAGAGAGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 889
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
752 CCCACTTATCTTCCATGATTTGCCAGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAG 811
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
890 CTTTCTACATGTCACAGACTGTGAGAGCATCTGCGGAGATTTTGTGTGAACCCAG 949
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
812 AAAAAAGCTAGCTGAGAGCAATTCAGAAACATCTTGG 852
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
950 CTAAGCGTGTACTCTGAGCAAAATCATGAAGCAAAATGG 990
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 14

AAFA4655
ID AAFA4655 standard; cDNA: 1594 BP.

AC AAFA4655;

DT 27-MAR-2001 (first entry)

DE Novel protein kinase cDNA, SEQ ID NO: 35.

XX Human; mouse; protein kinase; antirheumatic; antisclerotic; osteopathic;

XX immunosuppressive; cardiant; renal; antineoplastic; antistimatic;

XX dermatological; antidiabetic; antineoplastic; gene therapy; vaccine;

XX cancer; autoimmune disorder; cardiovascular disease; neurodegenerative disease;

XX inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.

OS Mus musculus.

PN WO200073469-A2.

PD 07-DEC-2000.

PF 26-MAY-2000; 2000MO-US14842.

PR 28-MAY-1999; 99US-0136503.

PA (SUGEN) SUGEN INC.

PI Plowman GD, Martinez R, Whyte D, Sudersanam S;

DR WPT: 2001-032161/04.

PS P-PSDB: AAB65629.

XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and

XX treating immune-related diseases and disorders, cardiovascular disease,

XX neurodegenerative diseases and/or cancers.

XX Disclosure: Fig 2; 310pp; English.

XX The present sequence encodes a novel protein kinase. The nucleic acids

XX and the protein kinases they encode may be used in the treatment and

XX diagnosis of diseases associated with inappropriate kinase expression

XX such as immune-related diseases and disorders, cardiovascular disease,

XX neurodegenerative diseases and/or cancers. The nucleic acids and

XX complementary sequences may also be used as DNA probes in diagnostic

XX assays. The kinase polypeptides may be used as antigens in the production

XX of antibodies of kinase expression and activity. Anti-kinase antibodies

XX and kinase antagonists may also be used to down regulate kinase

XX expression and activity. Diseases related to kinase expression and

XX activity include rheumatoid arthritis, atherosclerosis, autoimmune

XX disorders, complications of organ transplantation, myocardial infarction,

XX oxidative-stress related disorders, strokes, renal failure,

XX chronic inflammatory pelvic disease, multiple sclerosis, asthma,

XX osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and

XX Sequence 1594 BP; 376 A; 488 C; 458 G; 272 T; 0 other;

Query Match 11.2%; Score 261.6; DB 22; Length 1594;

Best Local Similarity 58.9%; Pred. No. 3.9e-47;

Matches 450; Conservative 0; Mismatches 314; Indels 0; Gaps 0;

```

89 ATGTGGCCCTATGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
80 ATGTGGGCACTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
149 TGGGGTCCATGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 208
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
140 TGGCTGCGCATATCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 199
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
209 TGTGGAGTGTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 268
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 200 TGAACCCAGTAGCTGCAGAGCTGTTCAAGAGAGTCCGAATTTATGAAGGACCTCAAC 259
 QY 269 ACCGACATGTCCTCAACCTCCAGACGCTTACAGAGAACAGAAATATTTGACCTGGTTC 328
 Db 260 ACCGACATGTCCTCAACCTTTTGTAGATGATAGACAGCAAGACCTATCTGCTGTA 319
 QY 329 TGGAGCAGCTCGGGGGGTGAGCTATTCGACTACCTGCTGTAAGAGAGGAGACCTGACGC 388
 Db 320 TGGAAATACGCTAGCGGAGAGAGTGTGACTACCTGCTGCTGCGACGCGCATGAAGG 379
 QY 389 CCAAGAGGCGCGAAAGTTCTTCCGCCAGATGTGTGCTGCGCTGACTTTCGACACACT 448
 Db 380 AGAAGAGGCTCGAGCAAGTTCGGGAGATGCTGACGCGGTGACACTGTCTATGACA 439
 QY 449 ACTCCATCTCCACAGACGCTTAAAGCCGAGAACCTGCTTTGGATGAGAAAAACA 508
 Db 440 AGAACATTTGACACAGGATCTTAAAGGCTGAAACCTGCTGCTGCTGCGGACCA 499
 QY 509 TCCGATTTGAGAGCTTGGGATGCGCTCCCTGACAGTGGGGGAGACCTCCTCGAGACCA 568
 Db 500 TCAAAATCGCCGACTTGGCTTTCAGCAATGATTCACGCTGGGCTCCAGCTGGACACT 559
 QY 569 GCTGCGGCTCCCGCATATGCTGCTCCAGAGTGAATTAAGGGGAAATATGATGGCC 628
 Db 560 TCTGTGGAGACCCCGCATAGCGCGCCGACAGGCTGTTCCAGGCAAGATATGATGGCC 619
 QY 629 GCGGCGACAGTGTGAGCTGTGATCATCTCTTCCCTGCTGCTGCGGGGCTGTC 688
 Db 620 CAGAGTGTGACATGTGAGGCTGGTGTCTGATCCCTGATACCTGCTGACGCGCTCCCTGC 679
 QY 689 CCTTGTATGACGACACCTCCGCGAGTGTGCTGAGAGGTAAGGAGGAGCTCTCCACA 748
 Db 680 CCTTCATGGGACACCACTTCAAGAGCTGCGGAGCGAGTCTTCAAGAGAAATACCGGG 739
 QY 749 TGGCCCATCTTCTCTCAAGATTGCCAGGCTCTGAGGGAATGATCGAAGTGAGC 808
 Db 740 TCCCTTCTCATGCTGACAGCTGCGAGAGCATTTGCGAGATTTCTGCTGTAAC 799
 QY 809 CCGAAAAAGGCTCAGCTGAGCAATTCAGAAACATCTTGG 852
 Db 800 CCGCAAAACGCTTACTGTGAGCAATTCATGAAGCAAAATGG 843

RESULT 15

ABA05737 standard; cDNA; 3170 BP.

ABA05737;

04-MAR-2002 (first entry)

Mouse; neuronal serine threonine protein kinase 9B5 coding sequence.
 Mouse; neuronal serine threonine protein kinase; apoplexy; cytosolic;
 cerebrioprotective; neuroprotective; antiparkinsonian; immunosuppressive;
 virulence; anti-HIV; antileukemic; antiinflammatory; antitumor; cancer;
 cerebral infarction; neurological disease; immunological disease; atopy;
 infection; leukaemia; polyarthritides; Crohn's disease; ulcerative colitis;
 9B5; ss.

Mus musculus.

Key Location/Qualifiers

1..2175

/tag= a

/product= "9B5"

polyA_signal

3094..3099

/tag= b

W020018108-A1.

22-NOV-2001.

XX

XX

XX

XX

XX

PF 17-MAY-2001; 2001MO-EP05660.

PR 17-MAY-2000; 2000DE-1024171.

PA (BADI) BASF-LYNX BIOSCIENCE AG.

PI Schneider A, Klausner B, Fischer A, Newrzella D, Goetz B;
 PI Rosner M, Eisenhardt G, Kuer R, Trutzel A, Kammandel B;
 PI Jomana Naim S, Schwaninger M;

DR MPI; 2002-055696/07.

P-PSDB; ABB04431.

New polynucleotide, useful for the diagnosis, treatment and prevention
 of e.g. apoplexy, tumors and autoimmune disease, comprises a
 polynucleotide encoding the neuronal protein kinase

Claim 1; Page 52-53; 75pp; German.

The present invention provides the protein and coding sequences of two
 neuronal serine threonine protein kinases each from the mouse and human.
 These are designated 9B5 and 9B5.b. Detection of their expression levels
 can be used to diagnose the risk of apoplexy and in prognosis of cerebral
 infarction. They can also be used to identify agents useful in the
 prevention, treatment and diagnosis of neurological diseases, including
 stroke, multiple sclerosis, Parkinson's disease, amyotrophic lateral
 sclerosis, tumours, especially carcinoma, immunological disorders,
 including autoimmune diseases, atopy, viral (including human immune
 deficiency virus) infections, acute or chronic leukocytic or myeloid
 leukaemia, primary chronic polyarthritides, Crohn's disease and ulcerative
 colitis. The present sequence is the murine 9B5 coding sequence.

Sequence 3170 BP; 714 A; 988 C; 897 G; 571 T; 0 other;

Query Match 11.1%; Score 259.4; DB 24; Length 3170;

Best Local Similarity 57.4%; Pred. No. 1.4e-46; Matches 467; Conservative 0; Mismatches 346; Indels 0; Gaps 0;

QY 89 ATGTGGGCCCCCTATGCGCTGAGAGAGCGCTGGCAAGGACAGACAGCGCTGTTAAAC 148
 Db 80 ATGTGGGCAACTATGAGCTGCTGTAAGACATCGGAGAGGCAACTGCGCAAAATGCAAC 139
 QY 149 TCGGGTCCACTGATGAGGAGGTCAGAAAGTCCGCAATCAAGATGTGACCGGAGAAC 208
 Db 140 TGGCTGGCAATATCTACAGCGCGCGAGGTCGCTATTAGATCATGATTAAGACCCAGC 199
 QY 209 TGTGAGGTCGCTGATGATGAGGAGGAGATCGCATCTGAGCTCAATGCAAC 268
 Db 200 TGAACCCAGTACTGTGACAGAACTGTTCAGAGAAAGTCCGAATTTATGAAGGACTCAAC 259
 QY 269 ACCGACATGCTCTCAAGCTCCAGACGCTTACGAGAAACAAGAAATATTTGATCTGTTTC 328
 Db 260 ACCGCAACATCGTGAAGCTTTTGTAGGTATGAGAGCGAGAAAGACCTATACCTGGTGA 319
 QY 329 TGGAGCAGCTCTCGGGGGGAGCTATTCGACTACCTGTTAAAGAGGAGGAGTGAAC 388
 Db 320 TGGAAATACGCTAGCGGAGGAGAGTGTGACTACCTGCTGCGACAGCGCGCATGAAG 379
 QY 389 CCAAGAGGCGCGAAAGTTCTTCCGCCAGATGTGTGCTGCGAGCTTTCGCCACAGCT 448
 Db 380 AGAAGAGGCTCGAGCAAGTTCGGGAGATGCTGACGCGGTGACACTGCTGATGACA 439
 QY 449 ACTCCATCTCCACAGACGCTTAAAGCCGAGAACCTGCTTTGGATGAGAAAAACA 508
 Db 440 AGAACATTTGACACAGGATCTTAAAGGCTGAAACCTGCTGCTGAGAGCGAGCAACA 499
 QY 509 TCCGATTTGAGAGCTTGGGATGCGCTCCGAGTGGGGGAGACGCTCCTGAGACCA 568
 Db 500 TCAAAATCGCCGACTTGGGCTTACGACATGATGATTCACGCTGGGCTCCTAAGCTGACACT 559
 QY 569 GCTGCGGCTCCCGCATTTATGCTGTCACAGAGTGAATTAAGGGGAGAAATATGATGGCC 628
 Db 560 TCTGTGGAGAGCCCGCATAGCGCGCCGACAGAGCTGTTCCAGGGGAGAAAGATATGAGGC 619

QY 629 GCCGGCAGACATGTGGAGCTGTGAGTCATCTTCGCCCTGCTGCGGGCTTGC 688
DB 620 CAGAGGTGACATCTGGAGCTGGGTGTCACTCTGACAGCTGTGTCAGCGGCTCCCTGC 679
QY 689 CCTTGTATGACGAACTCCGCCAGCTGCTGGAGAGGTGAACGGGGCTTCCACA 748
DB 680 CCTTCGATGGGCAACCTCAGAGAGCTGGGAGCAATCTCAGAGGAAGTACCGGG 739
QY 749 TGGCCCACTTCATTCCTCCAGATTGCCAGAGCCCTCTGAGGGGAATGATCGAGTGGAGC 808
DB 740 TCCCTTCTACATGTCTACAGACTGCGAGAGCATTTCTGCGAGATTCTGTGCTGAAC 799
QY 809 CCGAAAAAGGCTCAGTCTGAGCAATTCAGAAACATCTTGGTACCTAGCGGGAAAC 868
DB 800 CCGCAAAAGCTGTACTCTGAGCAATTCAGAAACATGATGATCAACATCGGCTATG 859
QY 869 ACGAGCCAGACCCGCTGCTGAGAGCCAGCCCTG 901
DB 860 AGGGTGAGAGCTGAAGCCATACAGGAGGCTG 892

Search completed: April 21, 2003, 20:31:33
Job time : 337.629 secs

GenCore version 5.1.4-P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 19:25:40 ; Search time 3982.82 Seconds

(without alignments)
17054.732 Million cell updates/sec

Title: US-10-003-690-3

Perfect score: 2334
Sequence: 1 atgtctccggggccaaga.....ccaagggaccctctgccc 2334

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rod:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hlg_hum:*
40: em_hlg_mus:*
41: em_hlg_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2334	100.0	3109	AF479827	AF479827 Homo sapi
2	2256	96.7	2385	AX166526	AX166526 Sequence
3	2256	96.7	3007	AF479826	AF479826 Homo sapi
4	2223	95.2	2720	HSMB05307	AF479826 Homo sapi
5	2146	91.9	2576	AB058714	AB058714 Homo sapi
6	1470.6	63.0	2128	BC016681	BC016681 Homo sapi
7	961	41.2	2025	AX327995	AX327995 Sequence
8	961	41.2	2219	AX327993	AX327993 Sequence
9	818	35.0	1956	HSAB6701	AX327993 Sequence
10	538	23.1	1014	HSMB03233	AX327993 Sequence
11	491.8	21.1	3156	AK074411	AL831945 Homo sapi
12	446	19.1	3933	AB014885	AK074411 Homo sapi
13	391.2	16.8	3059	AF316542	AB014885 Halocynthia
14	373.6	16.0	5609	AY060288	AF316542 Caenorhab
15	371	15.9	38000	AC008974	AY060288 Drosophila
16	371	15.9	134792	AC020922	AC008974 Homo sapi
17	315.2	13.5	22606	AC079583	AC020922 Homo sapi
18	314	13.5	210105	AC125948	AC125948 Rattus no
19	307.2	13.2	1873	BC024291	BC024291 Homo sapi
20	265.8	12.7	2333	AF020089	AF020089 Homo sapi
21	265.8	11.4	3529	AB088047	AB088047 Homo sapi
22	265.8	11.4	3609	AB049127	AB049127 Homo sapi
23	264.2	11.3	2462	AX399987	AB049127 Homo sapi
24	264.2	11.3	3226	AY057448	AX399987 Sequence
25	264.2	11.3	3312	AX305105	AY057448 Homo sapi
26	264.2	11.3	3392	AX305106	AX305105 Sequence
27	264.2	11.3	4917	AB058763	AX305106 Sequence
28	261.6	11.2	1594	AX056390	AB058763 Homo sapi
29	259.4	11.1	3170	AX305103	AX056390 Sequence
30	259.4	11.1	3250	AX305104	AX305103 Sequence
31	245	10.5	2650	AB020480	AX305104 Sequence
32	245	10.5	4092	AF106937	AB020480 Rattus no
33	240.6	10.3	2112	AX375128	AF106937 Rattus no
34	240.6	10.3	2112	AF387638	AX375128 Sequence
35	240.6	10.3	2222	AX375131	AF387638 Homo sapi
36	240.6	10.3	2701	BC008771	AX375131 Sequence
37	240.6	10.3	2946	9 HSSPKEWK	BC008771 Homo sapi
38	238.4	10.2	2572	10 RNMAR2	X97630 H. sapiens m
39	237.2	10.2	1749	3 AF020310	283869 R. norvegicu
40	228.8	9.8	2763	10 MKEM	AF020310 Drosophila
41	227.8	9.8	2681	3 AF181649	X70764 M. musculus
42	225.4	9.7	1869	8 ATRKIN2	AF181649 Drosophila
43	225.4	9.7	2209	8 AY093170	X94757 A. thaliana
44	224.2	9.6	4498	10 MM011494	AY093170 Arabidops
45	221.4	9.5	2352	6 AX262512	U11494 Mus musculu

ALIGNMENTS

RESULT 1
LOCUS AF479827 3109 bp mRNA linear PRI 12-MAR-2002
DEFINITION Homo sapiens protein kinase-like protein mRNA, complete cds.
ACCESSION AF479827
VERSION AF479827.1 GI:19401873
KEYWORDS

SOURCE

ORGANISM Homo sapiens.
Homo sapiens

REFERENCE 1 (bases 1 to 3109)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
She,X.Y., Yu,L. and Guo,J.H.

TITLE Direct Submission
JOURNAL Submitted (02-FEB-2002) School of Life Sciences, Laboratory of

Pred. No. is the number of results predicted by chance to have a

Db	1958	CTGGGCTCCCTCCGTTTCACCGGGCAAGATGACGCTCCCTACCGCTGAGGAGATGCC	2017
QY	1741	AGCTTGAGCCAGAGTGTCTCTCCCGGAGGTGGCAAAAGCTCTGGTTCGGCAACTTCATC	1800
Db	2018	AGCTTGAGCCAGAGTGTCTCTCCCGGAGGTGGCAAAAGCTCTGGTTCGGCAACTTCATC	2077
QY	1801	TCCTTGACAAGAAAGAAACAATATTCCTCTGTCTAAAGAGCAAACTCTCAGCAGCATC	1860
Db	2078	TCCTTGACAAGAAAGAAACAATATTCCTCTGTCTGTGTATCCCAAGCTGAGTCAAGATGCTGTCA	2137
QY	1861	AAAGCAGCATCTGTCATGCTCTTCTGTGATTCGCCAGCTGTGAGTCAAGATGCTGTCA	1920
Db	2138	AAAGCAGCATCTGTCATGCTCTTCTGTGATTCGCCAGCTGTGAGTCAAGATGCTGTCA	2197
QY	1921	CAGACACACTTCAGGGCCGAGTACAAAGGCGAGTGGCGGCCCTCCGTCTCCAAAGGCC	1980
Db	2198	CAGACACACTTCAGGGCCGAGTACAAAGGCGAGTGGCGGCCCTCCGTCTCCAAAGGCC	2257
QY	1981	GTCGGCTTCAGAGTGGAGCATCAGCTCTCTGAGGGTCCAGAGCCCTCCCGGAGGGAGC	2040
Db	2258	GTCGGCTTCAGAGTGGAGCATCAGCTCTCTGAGGGTCCAGAGCCCTCCCGGAGGGAGC	2317
QY	2041	GGCAGCGGAGAGTGTGGGAGTCTACTCTCCGTCAACCTTCATCTATCTGGGTCCAGCCGT	2100
Db	2318	GGCAGCGGAGAGTGTGGGAGTCTACTCTCCGTCAACCTTCATCTATCTGGGTCCAGCCGT	2377
QY	2101	CGGTTCAAGCCAGTGGTGGAGACATCTCCAGCAGCAGCTCTGTGAGCAGTCATGACAGGCC	2160
Db	2378	CGGTTCAAGCCAGTGGTGGAGACATCTCCAGCAGCAGCTCTGTGAGCAGTCATGACAGGCC	2437
QY	2161	TCCTGTCAGAGCCCTGTGGCAGAGAGAAACGGGGCCAGACCCCGGCTGTGGTCCCA	2220
Db	2438	TCCTGTCAGAGCCCTGTGGCAGAGAGAAACGGGGCCAGACCCCGGCTGTGGTCCCA	2497
QY	2221	CCCGGAAGCTGTGAGCCGCCACCGGGCCGCCAGACCCAGAGCTGAGGAGCTCTCCCGC	2280
Db	2498	CCCGGAAGCTGTGAGCCGCCACCGGGCCGCCAGACCCAGAGCTGAGGAGCTCTCCCGC	2557
QY	2281	CGAGGCCCCCCCCAAGAGACAGAGCTCTGTGGCCACCAAGGGACCCCTGTGCC	2334
Db	2558	CGAGGCCCCCCCCAAGAGACAGAGCTCTGTGGCCACCAAGGGACCCCTGTGCC	2611
RESULT 2			
AX16526			
LOCUS	AX16526	2385 bp	DNA
DEFINITION	Sequence 17 from Patent WO0138503.		linear
ACCESSION	AX16526		
VERSION	AX16526.1	GI:14546871	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Ciriatiata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Plowman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R., Flanagan,P. and Clarry,D.S.		
TITLE	Novel human protein kinases and protein kinase-like enzymes		
JOURNAL	Patent: WO 0138503-A 17 31-MAY-2001.		
FEATURES	Sugen, Inc. (US)		
source	Location/Qualifiers		
	1..2385		
BASE COUNT	494 a 768 c 731 g 392 t		
ORIGIN			
	/organism="Homo sapiens" /db_xref="taxon:9606"		
Query Match	96.7%;	Score 2256;	DB 6;
Best Local Similarity	100.0%;	Pred. No. 0;	Length 2385;
Matches 2256;	Conservative 0;	Mismatches 0;	Indels 0;
			Gaps 0;
79	CACGCCCAATATGTGGCCCTATCGCTGGAGAAAGACGCTGGCAAGAGACAGACAGG	138	

Db	127	CAAGCCCAATATGTGGGGCCCTATCGGCTGTGAGAGAGCGTGGGCAAAAGGACAGACAGGG	186
QY	139	CTGCGTTAAACTCGGGGTCCACTGTGATACCGGGTTCAGAGAGTCCGCATCAAGATCGTAAC	198
Db	187	CTGGTTAAACTCGGGGTCCACTGTGATACCGGGTTCAGAGAGTCCGCATCAAGATCGTAAC	246
QY	199	CGGAGAACGCTGTGCGAGTGGTGTGTGATGAAGGTGAGGGGGGATCGGCATCTGTAG	258
Db	247	CGGAGAACGCTGTGCGAGTGGTGTGTGATGAAGGTGAGGGGGGATCGGCATCTGTAG	306
QY	259	CTCATGAAACCCACATGTCCTCAAGCTTCACAGCTCTACGAGAACGAATATTTG	318
Db	307	CTCATGAAACCCACATGTCCTCAAGCTTCACAGCTCTACGAGAACGAATATTTG	366
QY	319	TACCTGGTTCTGGAGACGCTCTGGGGGGTGACTATTCGACTTACCTGTGTAAAGAGGG	378
Db	367	TACCTGGTTCTGGAGACGCTCTGGGGGGTGACTATTCGACTTACCTGTGTAAAGAGGG	426
QY	379	AGACTACGCCCCAAGAGAGGCCCAAAAGTTCTCCGACGATTTGTGTGGCTGTGATTC	438
Db	427	AGACTACGCCCCAAGAGAGGCCCAAAAGTTCTCCGACGATTTGTGTGGCTGTGATTC	486
QY	439	TGCGACAGCTACTCCATCTGTCCACAGACCTTAAAGCCCGAAGACTGTCTTTGGATGAG	498
Db	487	TGCGACAGCTACTCCATCTGTCCACAGACCTTAAAGCCCGAAGACTGTCTTTGGATGAG	546
QY	499	AAAAACAACATCCGCATTTGAGACTTCGGCATGGATGGCCCTCGCAGGCGGGGGACAGCTC	558
Db	547	AAAAACAACATCCGCATTTGAGACTTCGGCATGGATGGCCCTCGCAGGCGGGGGACAGCTC	606
QY	559	CTGAGACACGCTGCGGGTCCCCCATTTATGCTGTCCAGAGTGTAAAGGGGAAAA	618
Db	607	CTGAGACACGCTGCGGGTCCCCCATTTATGCTGTCCAGAGTGTAAAGGGGAAAA	666
QY	619	TATGATGGCGCGGGGACAGCATGTGGAGCTGTGTGATCATTCCTTCCTCGCCTCTGTG	678
Db	667	TATGATGGCGCGGGGACAGCATGTGGAGCTGTGTGATCATTCCTTCCTCGCCTCTGTG	726
QY	679	GGGGCTGTGCTTTTGTATGACGACAACCTCCGCGAGTGTGTGAGAGGTGAACGGGGC	738
Db	727	GGGGCTGTGCTTTTGTATGACGACAACCTCCGCGAGTGTGTGAGAGGTGAACGGGGC	786
QY	739	GTCTTCACATGCCCCCATCTCATTTCTCCAGATTGCCAGAGCTCTGTGAGGGGATGATC	798
Db	787	GTCTTCACATGCCCCCATCTCATTTCTCCAGATTGCCAGAGCTCTGTGAGGGGATGATC	846
QY	799	GAAGTGAGCCCCAAAAAAGAGTCTGAGCTGAGCAATTTAGAAAAATCCTGTGTACTTA	858
Db	847	GAAGTGAGCCCCAAAAAAGAGTCTGAGCTGAGCAATTTAGAAAAATCCTGTGTACTTA	906
QY	859	GGCGGAAACACGAGCCGACGCCGTGCTGAGGACAGCCCTGTGGCGCGCGGGTAGCATG	918
Db	907	GGCGGAAACACGAGCCGACGCCGTGCTGAGGACAGCCCTGTGGCGCGCGGGTAGCATG	966
QY	919	CGGAGCTGTGCATCCAAAGGAGGTGTGAGCCCGAGCTCTTAAGAGACATGTGCATCTGTG	978
Db	967	CGGAGCTGTGCATCCAAAGGAGGTGTGAGCCCGAGCTCTTAAGAGACATGTGCATCTGTG	1026
QY	979	GGCTGCTTCAAGGACCGCGAGAGGTGTGCATCGGAGCTGTGCGAGTGAAGAGGAACAACAA	1038
Db	1027	GGCTGCTTCAAGGACCGCGAGAGGTGTGCATCGGAGCTGTGCGAGTGAAGAGGAACAACAA	1086
QY	1039	GAAAAATGATATATTTATCTGCTTTTGGATTCGGAAGAGCGGTATCCACGCTGTGAGAC	1098
Db	1087	GAAAAATGATATATTTATCTGCTTTTGGATTCGGAAGAGCGGTATCCACGCTGTGAGAC	1146
QY	1099	CAGGACCTGCTCCCGGGAATGATGTGACCCCCCGGAAGCTGTGGAATTCCTCATG	1158
Db	1147	CAGGACCTGCTCCCGGGAATGATGTGGAATTCCTCATG	1206
QY	1159	CTGAGCGTTCACGGAAAGGGGACCAAGAGCGGAAGTCTATGGAAGTCTTGACATCAC	1218

489 TTGATGAGAAAAACACATCCGATTCGAGACTTGGCATGGCGTCCCTGCAGGTGG 548
 301 TTGATGAGAAAAACACATCCGATTCGAGACTTGGCATGGCGTCCCTGCAGGTGG 360
 549 GGCAGACCTCTGAGAGACAGCTGGGGTCCCGCATTTATGCGTCCAGAGGTATTA 608
 361 GGCAGACCTCTGAGAGACAGCTGGGGTCCCGCATTTATGCGTCCAGAGGTATTA 420
 609 GGGGAAAAATATGATGGCGCGGAGACATGATGGAGCTGTGGAGTCACTCTTGGC 668
 421 GGGGAAAAATATGATGGCGCGGAGACATGATGGAGCTGTGGAGTCACTCTTGGC 480
 669 CCGCTGCTGGGGGCTTGGCTTTGATGAGACAACTCCCGCAGCTGTGGAGAGT 728
 481 CCGCTGCTGGGGGCTTGGCTTTGATGAGACAACTCCCGCAGCTGTGGAGAGT 540
 729 GAAACGGGCGTCTCCACATGCCCCACATTCCTCCAGATTCAGAGCTCTGAG 788
 541 GAAACGGGCGTCTCCACATGCCCCACATTCCTCCAGATTCAGAGCTCTGAG 600
 789 GGAATGATGAGATGAGCGCCGAAAAAGGCTGATGAGCAATTCAGAAATCC 848
 601 GGAATGATGAGATGAGCGCCGAAAAAGGCTGATGAGCAATTCAGAAATCC 660
 849 TTGCTACCTAGGGGAGAAACAGAGACCGCTGCTGGAGCGCCCTGGCGCG 908
 661 TTGCTACCTAGGGGAGAAACAGAGACCGCTGCTGGAGCGCCCTGGCGCG 720
 909 GGTAGCCTAGCGAGCTGCCATCCAAACGAGAGCTGGAGCCCGCAGTCTAGAGCAT 968
 721 GGTAGCCTAGCGAGCTGCCATCCAAACGAGAGCTGGAGCCCGCAGTCTAGAGCAT 780
 969 GGCATCATCTGGGCTGCTTCAGGGAGCGGAGAGGCTGATGCGGAGCTGGAGAG 1028
 781 GGCATCATCTGGGCTGCTTCAGGGAGCGGAGAGGCTGATGCGGAGCTGGAGAG 840
 1029 GGAACCAAGAAAAAGATATATATATCTCTTTGGATGGAGAGAGCGGTATCCAG 1088
 841 GGAACCAAGAAAAAGATATATATATCTCTTTGGATGGAGAGAGCGGTATCCAG 900
 1089 CTGTGAGAGACAGGACCTGCTCCCGGAAATGATGTTGACCCCGCGGAAAGCTGGA 1148
 901 CTGTGAGAGACAGGACCTGCTCCCGGAAATGATGTTGACCCCGCGGAAAGCTGGA 960
 1149 TTCTCCATGCTGAGCGCTACGGGAAAGCGGACAGAGAGGAGTCAATGAGTCT 1208
 961 TTCTCCATGCTGAGCGCTACGGGAAAGCGGACAGAGAGGAGTCAATGAGTCT 1020
 1209 GAGCATACCGGAGTCCGGGGTGTGCTCCCTGTACCAACCGGAGCGGCTTGGAGAT 1268
 1021 GAGCATACCGGAGTCCGGGGTGTGCTCCCTGTACCAACCGGAGCGGCTTGGAGAT 1080
 1269 GGGCCAGACAGGACAGATCCCTAGCGTATGAGAGCTCCAGAGGCTCTGCTCCAG 1328
 1081 GGGCCAGACAGGACAGATCCCTAGCGTATGAGAGCTCCAGAGGCTCTGCTCCAG 1140
 1329 CCGCTTAAGACAGCCAGAGAGTCCGCTTTCTCTTCCACCGAGCGGGGGCTGAGA 1388
 1141 CCGCTTAAGACAGCCAGAGAGTCCGCTTTCTCTTCCACCGAGCGGGGGCTGAGA 1200
 1389 TGAAGCTGAGAGCGGGGCTCCCGACTTCCAAAACGAGAGCGTCCCTCTCGGGGCC 1448
 1201 TGAAGCTGAGAGCGGGGCTCCCGACTTCCAAAACGAGAGCGTCCCTCTCGGGGCC 1260
 1449 CAGGCGTGGGGGCGCGGGGAGAGCGCCCGCCCGCAGTCCCGCTCCACACCCCTGCG 1508
 1261 CAGGCGTGGGGGCGCGGGGAGAGCGCCCGCCCGCAGTCCCGCTCCACACCCCTGCG 1320
 1509 CAGGCGTGGGGGCTCCCGGCTCTCTGAGGGAGAGCGCTTGCAGTGGCTCTGCACAC 1568
 1321 CAGGCGTGGGGGCTCCCGGCTCTCTGAGGGAGAGCGCTTGCAGTGGCTCTGCACAC 1380

1569 GCCCGGGGCGAGTCCACCGGAGACCCCGGGAGACACACACCCCGGCGGGTGG 1628
 1381 GCCCGGGGCGAGTCCACCGGAGACCCCGGGAGACACACACCCCGGCGGGTGG 1440
 1629 CGTGGGGGAGCGCGCTGGAGAGTCTCTCACTCCATCCGACAGCTTCTGGGCTC 1868
 1441 CGTGGGGGAGCGCGCTGGAGAGTCTCTCACTCCATCCGACAGCTTCTGGGCTC 1500
 1689 CCGTCTTTACCGGGGCAAGATGAGAGTCCCTACCGCTGAGAGAGTCCAGCTTGC 1748
 1501 CCGTCTTTACCGGGGCAAGATGAGAGTCCCTACCGCTGAGAGAGTCCAGCTTGC 1560
 1749 GCCAGAGTCTCCCGGAGAGTGGCAAAACCTCTGCTGGGAACTTCATCTCTTGA 1808
 1561 GCCAGAGTCTCCCGGAGAGTGGCAAAACCTCTGCTGGGAACTTCATCTCTTGA 1620
 1809 CAAAGAAAGAAATATCTCTGCTGCTTAAAGAGACAACTCTCAGCAGCATCAAGAGA 1868
 1621 CAAAGAAAGAAATATCTCTGCTGCTTAAAGAGACAACTCTCAGCAGCATCAAGAGA 1680
 1869 CATGCTCATGCTCTCTCTGATGATCCCGAGCTGAGACAGTGTGCTCAGAGACAG 1928
 1681 CATGCTCATGCTCTCTCTGATGATCCCGAGCTGAGACAGTGTGCTCAGAGACAG 1740
 1929 CTTACAGGGCGAGTACAGAGGCGAGTGGCGGCGCTCTGCTTCCAAAAGCCGCTGCT 1988
 1741 CTTACAGGGCGAGTACAGAGGCGAGTGGCGGCGCTCTGCTTCCAAAAGCCGCTGCT 1800
 1989 CCAGGTGAGATCAGTCTCTGAGAGTCCAGAGCTCTCCCGGAGAGGGAGCGAGGG 2048
 1801 CCAGGTGAGATCAGTCTCTGAGAGTCCAGAGCTCTCCCGGAGAGGGAGCGAGGG 1860
 2049 AGGTGATGATCTACTCTCTGATGATCCCTGATCTCTGATCTCTGATCTCTGATCT 2108
 1861 AGGTGATGATCTACTCTCTGATGATCCCTGATCTCTGATCTCTGATCTCTGATCT 1920
 2109 GCGAGTGTGAGACCATCCAGGACAGCTCTGAGAGCTCTGAGAGCTCTGAGAGCT 2168
 1921 GCGAGTGTGAGACCATCCAGGACAGCTCTGAGAGCTCTGAGAGCTCTGAGAGCT 1980
 2169 GGGCTGTGAG 2228
 1981 GGGCTGTGAG 2040
 2229 CCGTCAAGCCCGGAG 2288
 2041 CCGTCAAGCCCGGAG 2100
 2289 CCGCAAGGACAGAGAGTCTGCGCACCAAGGAGAGAGAGAGAGAGAGAGAGAGAG 2334
 2101 CCGCAAGGACAGAGAGTCTGCGCACCAAGGAGAGAGAGAGAGAGAGAGAGAGAG 2146

RESULT 5
 BC016681 2128 bp mRNA linear PRI 05-NOV-2001
 LOCUS Homo sapiens, clone MGC:21870 IMAGE:4341838, mRNA, complete cds.
 DEFINITION
 ACCESSION BC016681
 VERSION BC016681.1 GI:16741775
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 2128)
 AUTHORS Strausberg R.
 TITLE Direct Submission
 JOURNAL Submitted (31-OCT-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 CONTACT: MGC help desk

Email: cgraphs@email.nih.gov
 Tissue Procurement: Lou Staudt
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villalona@bcm.tmc.edu
 Villalón, D.K., Luna, R.A., Hale, S.M., Huiyk, S., Lu, X., Garcia,
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
 Muzny, D.M., Gibbs, R.A.

Clone distribution: MCC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRK Plate: 26 Row: 3 Column: 5
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES

source
 1. .2128
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MCC:21870 IMAGE:4341838"
 /tissue_type="Lymph, lymphoma"
 /clone_id="NH_MCC_85"
 /lab_host="DH10b"
 /note="Vector: pCMV-SPORT6"
 227. 1648
 /codon_start=-1
 /product="Unknown (protein for MCC:21870)"
 /protein_id="AAH1681.1"
 /db_xref="GI:1674176"
 /translation="MRSIPNGELDPVLESMAISAGCFRDRERLHRELSEENOEK
 IYLLDRKREYSCEDODLPKRVDPKRVSPMLSRGRPEKSEVSTID
 AGGSGPVPKRALMAHOSRSVSGATSSSPSRSPVSPSPGADPA
 RGGSPSTKQTLPSPKRGAGAGBPSPSPARSTPLDPPGSPSSSGTPLHSLHT
 PRASPTGPTGTPSPSPGAGVGAAMSRSLNSISPLSGSPRHRKKQVPLAEKMS
 LPESSPELAKRMFNFISLDEKEIFLVLDKPLSLKADIVHAFISISLSHSL
 SQTSPFAEYKASGSPVFKPVRFOVDISSEGEPSRRRSGSGGGSYVFTLSG
 PSRRKRVETIOALSTHDOPVALADEKNAGQRRPAPAPPSLQPPGRRPPEL
 SSSRRAPRDKRLIANNGTPLP"

CDS
 227. 1648
 /product="Unknown (protein for MCC:21870)"
 /protein_id="AAH1681.1"
 /db_xref="GI:1674176"
 /translation="MRSIPNGELDPVLESMAISAGCFRDRERLHRELSEENOEK
 IYLLDRKREYSCEDODLPKRVDPKRVSPMLSRGRPEKSEVSTID
 AGGSGPVPKRALMAHOSRSVSGATSSSPSRSPVSPSPGADPA
 RGGSPSTKQTLPSPKRGAGAGBPSPSPARSTPLDPPGSPSSSGTPLHSLHT
 PRASPTGPTGTPSPSPGAGVGAAMSRSLNSISPLSGSPRHRKKQVPLAEKMS
 LPESSPELAKRMFNFISLDEKEIFLVLDKPLSLKADIVHAFISISLSHSL
 SQTSPFAEYKASGSPVFKPVRFOVDISSEGEPSRRRSGSGGGSYVFTLSG
 PSRRKRVETIOALSTHDOPVALADEKNAGQRRPAPAPPSLQPPGRRPPEL
 SSSRRAPRDKRLIANNGTPLP"

BASE COUNT 408 a 707 c 651 g 362 t
 ORIGIN

Query Match 63.0%; Score 1470.6; DB 9; Length 2128;
 Best local Similarity 99.7%; Pred. No. 3.8e-256;
 Matches 1473; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

858 AGCGGGAAACAGACGACCCGCTGAGGACCCCTGGCCGGGGTAGGCAT 917
 169 AGCGGGAAACAGACGACCCGCTGAGGACCCCTGGCCGGGGTAGGCAT 228
 918 GCGGAGCTGCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 977
 229 GCGGAGCTGCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 288
 978 GCGGAGCTGCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1037
 289 GCGGAGCTGCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 348
 1038 AGAAAGATGATATATATCTGCTTTTGGATCGAAGAGAGAGAGAGAG 1097
 349 AGAAAGATGATATATATCTGCTTTTGGATCGAAGAGAGAGAGAGAG 408
 1098 CAGAGAGCTGCTCCCGGAGATGATGACCCCGCGGAAAGCTGAGATTC 1157
 409 CAGAGAGCTGCTCCCGGAGATGATGACCCCGCGGAAAGCTGAGATTC 468
 1158 GCTGAGCCGTCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1217
 469 GCTGAGCTGTCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 528

1218 CGATGCCGGGGGTGTGTGCTTCCCTGTACCCAGCCGAGCCCTTGAGATGCCACCA 1277
 529 CGATGCCGGGGGTGTGTGCTTCCCTGTACCCAGCCGAGCCCTTGAGATGCCACCA 588
 1278 CAGCAGAGATCCCGTAGCGGTAGTGGAGCTCCACGGGTGTGTCTCAGAGCCCTAG 1337
 589 CAGCAGAGATCCCGTAGCGGTAGTGGAGCTCCACGGGTGTGTCTCAGAGCCCTAG 648
 1338 CAGCCCAAGAGATCCCGTAGCGGTAGTGGAGCTCCACGGGTGTGTCTCAGAGCCCTAG 1397
 649 CAGCCCAAGAGATCCCGTAGCGGTAGTGGAGCTCCACGGGTGTGTCTCAGAGCCCTAG 708
 1398 AGCGGGGGGTCTCCCGACCTTCCAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 1457
 709 AGCGGGGGGTCTCCCGACCTTCCAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 768
 1458 GGGGGGGGGGAG 1517
 769 GGGGGGGGGGAG 828
 1518 AGGCTCCCGCGCTCTCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1577
 829 AGGCTCCCGCGCTCTCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 888
 1578 CAGTCCCGAG 1637
 889 CAGTCCCGAG 948
 1638 AGCGCGCTGAG 1697
 949 AGCGCGCTGAG 1708
 1698 TCACCGCGAG 1757
 1009 TCACCGCGAG 1808
 1758 CTCCCGGAG 1817
 1069 CTCCCGGAG 1828
 1818 ACAATTTCTCTGCTGCTTAAAGCAAACTCTGAGAGAGAGAGAGAGAGAGAGAGAG 1877
 1129 ACAATTTCTCTGCTGCTTAAAGCAAACTCTGAGAGAGAGAGAGAGAGAGAGAGAG 1888
 1878 TGCCCTTCTGCTGCTTAAAGCAAACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1937
 1189 TGCCCTTCTGCTGCTTAAAGCAAACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1248
 1938 CGAGTCAAG 1997
 1249 CGAGTCAAG 1308
 1998 CATCAGCTCTGAG 2057
 1309 CATCAGCTCTGAG 1368
 2058 CATCAGCTCTGAG 2117
 1369 CATCAGCTCTGAG 1428
 2118 GAGAGATCCAG 2177
 1429 GAGAGATCCAG 1488
 2178 GAGAGATCCAG 2237
 1489 GAGAGATCCAG 1548
 2238 CCAAGCCGGGGGAG 2297
 1549 CCAAGCCGGGGGAG 1608
 2298 CCAAGCCGGGGGAG 2334

```

Db 1609 CAGAGAGCTCTGGCCACCAAGGGAGCCCTCTGCC 1645
RESULT 7
AX327995
LOCUS AX327995 2025 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 3 from Patent W00181588.
ACCESSION AX327995
VERSION AX327995.1 GI:18098148
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1. Meyers, R.
2246. Protein Kinase molecules and uses therefor
PATENT: WO 0181588-A 3 01-NOV-2001.
JOURNAL Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1. 2025
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 442 a 641 c 616 g 326 t
ORIGIN
Query Match 41.2%; Score 961; DB 6; Length 2025;
Best Local Similarity 70.1%; Pred. No. 5.8e-164;
Matches 1477; Conservative 0; Mismatches 435; Indels 195; Gaps 6;
QY 74 CCACAGACCCCAATATGTGGCCCTATCGCTGAGAGAAACGCTGGCAAGACAGAGA 133
DB 29 CGAGACAGCGCGATATGTGGCCCTATCGCTGAGAGAAACGCTGGCAAGACAGAGA 88
QY 134 CAGGCGTGTAACTGGGGTCCATCGATCAGGCGTCAAGAGTCCGCAATCAATCG 193
DB 89 CAGCTCGGTGAAGCTGGGGTTCATCGCTCAGCTGCAAGAGTGGCCATCAAGATCG 148
QY 194 TGAACCGGAGAAAGCTGTGAGTGTCTGATGAAGAGTGGAGCGGAGATCGCATCC 253
DB 149 TCACCGGTGAAGAGCTGAGAGTGTCTGATGAAGAGTGGAGCGGAGATCGCATCC 208
QY 254 TGAAGTCATCGAAGACCCCAATGTCTCAAGCTCCACAGAGCTGACAGAGAAAT 313
DB 209 TGAAGTCATGTGAGACCCCAATGTCTCAAGCTCCACAGAGCTGACAGAGAAAT 268
QY 314 ATTTGATCTGTCTGTGAGACAGCTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAG 373
DB 269 ATTTGATCTGTCTGTGAGACAGCTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAG 328
QY 374 AGGAGAGACTGAGCGCCCAAGAGAGCGGAAAGTTCTTCCGCAAGTGTGTGCGCTGG 433
DB 329 AGGAGAGACTGAGCGCCCAAGAGAGCGGAAAGTTCTTCCGCAAGTGTGTGCGCTGG 388
QY 434 ACTTTCGCACTACTCTCATCTGCAAGAGACTAAACCCGAGAACTGCTTTGG 493
DB 389 ACTTTCGCACTACTCTCATCTGCAAGAGACTAAACCCGAGAACTGCTTTGG 448
QY 494 ATGAGAGAAACCAATCGCATGTGACAGCTTCCGATGAGCTGCTCCGAGGAGGAGCA 553
DB 449 ATGAGAGAAACCAATCGCATGTGACAGCTTCCGATGAGCTGCTCCGAGGAGGAGCA 508
QY 554 GCTCTGTGAGACAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 613
DB 509 GCTCTGTGAGACAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 568
QY 614 AAAAATATATGAGCGCGCGGAGAGATGTGAGAGCTGTGAGTGTGAGTGTGAGTGTGAG 673
DB 569 AGAAGATATGAGCGCGGAGAGATGTGAGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 628
QY 674 TGTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 733

```

```

Db 629 TGTGAGGAGCTGTGCTTGTGAGAGTGAACAATTGCAAGAGCTGTGAGAGTGAAC 688
QY 734 GGGGCTCTTCCACATGCCCCCACTTCAATCTCCAGATGTCAGAGAGCTCTGAGGGAA 793
DB 689 GGGGCTCTTCCACATGCCCCCACTTCAATCTCCAGATGTCAGAGAGCTCTGAGGGCA 748
QY 794 TGTGAGTGTGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 853
DB 749 TGTGAGTGTGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 808
QY 854 ACCTAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 913
DB 809 ATATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 862
QY 914 CCATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 973
DB 863 AGATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 922
QY 974 CACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1033
DB 923 CACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 982
QY 1034 ACCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1093
DB 983 ACCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1042
QY 1094 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1153
DB 1043 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1102
QY 1154 CCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1213
DB 1103 CCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1162
QY 1214 TCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1273
DB 1163 TCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1213
QY 1274 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1333
DB 1214 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1273
QY 1334 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1393
DB 1274 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1286
QY 1394 CTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1453
DB 1287 CTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1286
QY 1454 GTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1513
DB 1287 GTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1326
QY 1514 CCCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1573
DB 1327 CCCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1357
QY 1574 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1633
DB 1358 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1411
QY 1634 GGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1693
DB 1412 GGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1471
QY 1694 GCTTTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1753
DB 1472 GCTTTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1531
QY 1754 AGTCCTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1813
DB 1532 AGTCCTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1591

```

QY	1814	AAGAACAAATATTCCTCTGCTGTAAGGACCAACCCCTCAGACAGATCAAGACAGATCG	1873
Db	1592	AGGACAGATCTTGCTGCTGATCATCAAGACCAACCTCTAGCTCCATCAAGGCTGACATCG	1651
QY	1874	TCGATGCCCTTTCTGTGCATGCCAGCCAGCCCTGAGTCAACAGTGTGCTGTCCACAGACCACTTCA	1933
Db	1652	TGCACGCCCTTCTGTGCATGCCAGCCAGCCCTGAGTCAACAGTGTGCTGTCCACAGACCACTTCA	1711
QY	1394	GGGCGGAGTACAAAGCCAGTGGCGCCCTCTGCTTCCAAAGACCCGCTCCGCTTCCAGG	1993
Db	1712	GGGCGGAGTACAAAGCCAGTGGCGCCCTCTGCTTCCAAAGACCCGCTCCGCTTCCAGG	1771
QY	1994	TGACATCAGCTCCGCTGTAAGGCTCCAGAGCCCTCCGCGACGGGAGGGCAGCGGAGTG	2053
Db	1772	TGTATATCACTTACACAGAGGCT-----GGGAGGCGCAGAGAGGAGA	1813
QY	2054	GTGGCATTCTACTCCGTCACCTTTCATCTGCGGGTCCAGCCGCTGTTCAAGCAG	2113
Db	1814	ACGGCATCTACTCCGTCACCTTTCATCTGCGGGTCCAGCCGCTGTTCAAGAGG	1873
QY	2114	TGGTGAGAGCATCTCAGGACACAGCTCCTGAGACACTCATAGACAGCCCTCGCTGCAGGCC	2173
Db	1874	TGGTGAGAGCATCTCAGGACACAGCTCCTGAGACACTCATAGACAGCCCTCGCTGCAGGACT	1933
QY	2174	TGCGACA	2180
Db	1934	TGTCAGA	1940
RESULT 8			
LOCUS	AX327993	2219 bp	DNA
DEFINITION	Sequence 1 from Patent WO0181588.	Linear	PAT 07-JAN-2002
ACCESSION	AX327993		
VERSION	AX327993.1	GI:18098146	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	Meyers, R.		
TITLE	2246, protein kinase molecules and uses therefor		
JOURNAL	Patent: WO 0181588-A 1 01-NOV-2001; Millennium Pharmaceuticals, Inc. (US)		
FEATURES	Location/Qualifiers		
source	1..2219		
CDS	1..2219 /organism="Homo sapiens" /db_xref="taxon:9606" 53..2077 /note="unnamed protein product" /cdoe_start=-1 /protein_id="CAD20436.1" /db_xref="GI:18098147" /translation="MTSTGKDGGAHQAYVYPRLEKTLGKGQGTGLVKGVCYTCOK VAIKIVREKLSSEVLAKVEREIALILKLIHPYLKLDVYENKRYLYLVEHSGGEE LFDLYVKGRLTPREARKFROIISALDFCHSHICHRDLKPEVILLDEKNNTIADF GMAISYGDLSLSTSGSPHYACDEVTREGKYGSRKADWSCGTYTALIVGALPRDD DNAROLEKRCRYFHHPHPTPPCCSLGMSGVYDARRLITLHIDKHITJGCKNE PEPDPILEPKVQLKSLPDSLEDIDVDVDSHMSJCFPDNRNKLDDLISEENOKMITY FLIDRKERYPSQEDEDLPENNETIDPRKRVDSPLMNRHCRREPRKSMELSVTDGCG SPVAPARAIEMAQHGORSISGASGLSTSPSSPVTPHPSPRSGSLPTPKGTPIVH TPRESPLAGTNPPTPPSPSYGVGVMARLINSKFSFGSPREHRRKLOVTPPEMSML TPRESPLAKSMNGNFIISLEKEOIFVYIKDRPLSIKADIYVHATISISLSHSYS OTSPRAEYKATGGAAYQKRYQVDTTTEGGAGACQENCTISTYTPILISPSRFRKR VETIIQQLISTHDPPEAQLSEPPPPAPGLSMWAGLKGQVATVYESSL"		
BASE COUNT	476 a	716 c	683 g
ORIGIN	716 c	683 g	344 t

QY	74	CCGACAGCGCCAAATATGTGGGCCCGATGCGGTGGAGAAAGCGTGGCCAAAGGACAGA	133
Db	81	CGCAGCAGCGCCAAATATGTGGGCCCGATGCGGTGGAGAAAGCGTGGCCAAAGGACAGA	140
QY	134	CAGGCGTGTAAACTGCGGGGTCACCTGCATYCACGGGGTGAAGAGTCCGATCAAGATCG	193
Db	141	CAGGCTGTGTGAAGCTGGGGGTTCAGTGGGTACCTGTCCAGAAAGGTGGCATCAAGATCG	200
QY	194	TGAACGGGAGAAAGCTGTGGAAGTCCGATGTGATGAAGTGGAGCGGAGATCGCATCC	253
Db	201	TCAACCGCTGAAAGCTCAGCGAGTCCGGTGTGTATGAAGTGGAGCGGAGATCGCATCC	260
QY	254	TGAAGCTCATCGAAACACCGCATGTCTCCATGAGTCCAGCAAGCTCATCGAACAAGAAAT	313
Db	261	TGAAGCTCATTTGAGCACCCACCGATCTCTTAAGGTGACAGACCTTTATGTAAAAACAAAAT	320
QY	314	ATTGTATCTGTGTCTTGAGACACGTCTCGGGGGGTGAGCTATTTGCATCTACGTGTAAAGA	373
Db	321	ATTGTATCTGTGTCTTGAGAACACGTGTCAAGTGTGATGAGTCTTTCGATCTACGTGTAAAGA	380
QY	374	AGGGGAGACTGAGCGCCCAAGAGAGCCGAAAGTTCTTCCGCAGATTGTGTGCGCTGG	433
Db	381	AGGGGAGGCTGAGCCCTTAAGAGAGGCTCGAAGATTTCTTCGGAGAGATCATCTCGCGCTGG	440
QY	434	ACTTTGTGCAAGGTATCTCATCTGCGACAGAACCTTAAGGCCCGAGACGTCCTTTGG	493
Db	441	ACTTGTGTGCAAGGCTACTTCATATGCAACAGGATGTGAACCTGTAAACCTCTCTGCTGG	500
QY	494	ATGAGAAAAACAATCCGGATTGTCAGACTTCGGCATGCGTCCCTCGAGGTGGGGGACA	553
Db	501	ACGAGAAAGAACAACTCCGATCGAGACTTTTGGATGGGTGCTCCGAGTGTGGGACA	560
QY	554	GCCTCTCGAGACCACTGGGGGTCCCCCATTTATGCGTGTCAAGAGTGTATTAAGGGGG	613
Db	561	GCGCTGTGGAGACCACTGTGGGTCCCGCCACATACCTCTGCCCCGAGGTGATCCGAGGGG	620
QY	614	AAAATATATGATGGCGCGCGGAGACATGTGAGAGCTGTGAGAGTCAATCCCTTCGCGCTGC	673
Db	621	AGAAATATGACGGCGCGGAGAGCGAGGTGTGAGAGCTGCGCGCTCATCTGTTCGCTTTC	680
QY	674	TCGTGGGGGCTCTGCGCTTTGATGAGACAACCTCCGCAAGCTGCTGTGAGAAAGTGAAC	733
Db	681	TGCTGGGGGCTCTGCGCTTTGAGAGATGACAATCTTGGAACAGCTGTGAGAAAGTGAAC	740
QY	734	GGGGCGTCTTCCACATCCGCCCATCTTATCCCGCGCACTGCGAGACTGTGTACGGGGGA	793
Db	741	GGGGCGTGTCTCACATCCGCCCATCTTATCCCGCGCACTGCGAGACTGTGTACGGGGGA	800
QY	794	TGATTCGAAGTGGAGCGCCGAAAAAAGCGTCAAGTGTGAGCAAAATTGAGAAACATTCCTGGT	853
Db	801	TGACGAGAGTGGAGCGCGCGCAAGCGCGCTCAAGTGTAGACAAATTGAGAAACATATGTGT	860
QY	854	ACCTAAGCGGGGAACAAGAGCCGACAGACCCGTGCTGAGCGCCAGCCCTTGCGCGCGGTAG	913
Db	861	ATTATAGGGGGGCAAGAAAGAGACCCGAAAC-----AGAGAGGCCCATTTCTCCAGAGGTGC	914
QY	914	CCATAGCGGAGCGCTGCGCATTCAGAGGAGGTGAGACCGCAAGCGCTTATGAGAGCATGGCAT	973
Db	915	AGATTCGCTGCTGCCAGCGCTGAGAGCAATGCAACCCGACGCGTGTGACACATGCACT	974
QY	974	CAGTGGGCTGCTTACAGGAGCGCGAGAGGCTGCATCGCGAGTGTGCGCACTGTGAGAGAGA	1033
Db	975	CAGTGGGCTGCTTCCGAGAGCGCGAACAAAGTGTGTGAGAGCGTGTGTCGAGAGAGAGA	1034
QY	1034	ACCAAGAAAAAGATGATATATTTATCTGCTTTTGGATGGAAGAGCGGTATCCAGCTGTG	1093
Db	1035	ACCAAGGAAGATATTTACTTCTCTCTCTGACCGGAAAGAAAGTCACTCCGACAGAG	1094
QY	1094	AGGACAGAGACCTGCTCCCGGATATGATTTGATACCCCGGAGAGGCTGTGAGATTCTC	1155
Db	1095	AGGATGAGACCTTCCCGCGAAGCAAGATAGACCTTCCCGGAGAGGCTGTGAGACTTCC	1156

OY 497 AGAAAAACATCCGATTCGACACTTCGATGCGCTCCCTGAGTGGGGACAGCC 556
 DB 241 AGAAGAACATCCGATTCGACACTTCGATGCGCTCCCTGAGTGGGGACAGCC 300
 OY 557 TCTGGAGACACCTGCGGGTCCCTCATTTGCGGTGTCAGAGGTGATTAAGGGGAAA 616
 DB 301 TGTGGAGACACCTGCGGGTCCCTCATTTGCGGTGTCAGAGGTGATTAAGGGGAAA 360
 OY 617 AATATATGCGCGCGGAGACATGTGAGAGTGTGAGTATCTCTGCGCTCTGCTG 676
 DB 361 AGATGACGCGCGGAGAGCGAGCTGTGAGAGTGTGAGTATCTCTGCGCTCTGCTG 420
 OY 677 TGGGGGCTCTGCGCTTGTGATGAGACAACTCCCGAGCTCTGAGAGGTGAAGAGG 736
 DB 421 TGGGGGCTCTGCGCTTGTGATGAGACAACTCCCGAGCTCTGAGAGGTGAAGAGG 480
 OY 737 GCGTCTTCACATGCGCGCTCATTTCTTCAGATGTCAGAGTGTGAGAGGTGAAG 796
 DB 481 GCGTCTTCACATGCGCGCTCATTTCTTCAGATGTCAGAGTGTGAGAGGTGAAG 540
 OY 797 TCGAATGAGACCGCGGAAAAAGCTGAGTGTGAGACAAATTCAGAAACATCTTGTAC 856
 DB 541 GCGAGGTGAGACCGCGGAAAAAGCTGAGTGTGAGACAAATTCAGAAACATCTTGTAC 600
 OY 857 TGGGGGAGAAACAGAGCCAGACCCGCTGAGAGCCAGCCGCTGCGGGGTAGCCA 916
 DB 601 TGGGGGAGAAACAGAGCCAGACCCGCTGAGAGCCAGCCGCTGCGGGGTAGCCA 654
 OY 917 TGGGGAGACCTGCGCTCATTCAGAGAGTGTGAGACCGCTGAGAGAGTGTGAGTAC 976
 DB 655 TCGGCTGCGCTGCGCTCATTCAGAGAGTGTGAGACCGCTGAGAGAGTGTGAGTAC 714
 OY 977 TGGGGTCTTCAGAGACCGCGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1036
 DB 715 TGGGGTCTTCAGAGACCGCGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 774
 OY 1037 AGAAGAAAGATATATATATATCTGTTTGTGAGTGTGAGAGTGTGAGAGTGTGAG 1096
 DB 775 AGAAGAAAGATATATATATCTGTTTGTGAGTGTGAGAGTGTGAGAGTGTGAG 834
 OY 1097 ACCAGAGACCTGCGCTCATTCAGAGAGTGTGAGACCGCTGAGAGAGTGTGAGTAC 1156
 DB 835 ATGAGAGACCTGCGCTCATTCAGAGAGTGTGAGACCGCTGAGAGAGTGTGAGTAC 894
 OY 1157 TGTGAGCGCTGAGAGACCGCGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1216
 DB 895 TGTGAGCGCTGAGAGACCGCGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 954
 OY 1217 CGGATGCGCGGGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1276
 DB 955 C-----GGAAGGCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 1005
 OY 1277 ACAGCCAGAGATCCGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1336
 DB 1006 ACAGCCAGAGATCCGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1065
 OY 1337 GCAAGCCAGAGATCCGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1396
 DB 1066 GCAAGCCAGAGATCCGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1078
 OY 1397 GAGGGGGGGGTCTCCGACTTCCAAAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAG 1456
 DB 1079 ----- 1078
 OY 1457 GGGGGCGGGGAGAGACCG 1516
 DB 1079 -----GAGCGCTCAGCGCTCAGCGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAG 1115
 OY 1517 CAGGCTCCCGCGCTCTCTGCGGGGAGACCGCTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1576
 DB 1116 -----ACCGCGAGGGGAGACCGCTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1149
 OY 1577 CCAAGTCCCG 1636

DB 1150 AGAGCCCGGGTGCACGCGCCCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1203
 OY 1637 GAGCGCGCGGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1696
 DB 1204 GGGTGGCTGAGAGGGCGCGCTCACTTCATCAAGAACAGTGTGAGTGTGAGTGTGAG 1263
 OY 1697 TTCACCGCGCGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1756
 DB 1264 TCCACCGCGCGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1323
 OY 1757 CTTCCCGCGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1816
 DB 1324 CTTCCCGAGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1383
 OY 1817 AACAAATATCTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1876
 DB 1384 AGCAGATCTTCTGCTGATCAAGAACAACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1443
 OY 1877 ATGCTTCTGCTGATCCCGAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1936
 DB 1444 AGCGCTTCTGCTGATCCCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1503
 OY 1937 CCGAGTCAAGAGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1996
 DB 1504 CCGAGTCAAGAGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1563
 OY 1997 ACATGAGCTCTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2056
 DB 1564 ATATCACTTACAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1605
 OY 2057 GCATCTACTGCTGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 2116
 DB 1606 GCATCTACTGCTGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1665
 OY 2117 TGAAGACATCCAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2176
 DB 1666 TGAAGACATCCAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1725
 OY 2177 CAGA 2180
 DB 1726 CAGA 1729

RESULT 10
 HSNB03233 1014 bp mRNA linear PRI 12-JUL-2002
 LOCUS
 DEFINITION Homo sapiens mRNA, cDNA DKFP547F069 (from clone DKFP547F069).
 ACCESSION AL831945
 VERSION AL831945.1 GI:21732471
 KEYWORDS
 SOURCE
 ORGANISM
 human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1014)
 Bloecker, H., Boecker, M., Brandt, P., Mewes, H.W., Well, B. and
 Wiemann, S.
 Direct Submission
 Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
 sequenced by GBF (National Research Centre for Biotechnology Ltd.,
 Braunschweig/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 This clone (DKFP547F069) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at <http://mips.gsf.de/proj/cdna/>.
 Location/Qualifiers
 FEATURES
 source
 1..1014
 /organism="Homo sapiens"


```

      /db_xref="taxon:9606"
      /chromosome="19"
      /clone="DKFZp547F069"
      /tissue_type="brain"
      /clone_lib="547 (synonym: hfb1). Vector pSport1; host
      DH10B; sites NotI + SalI"
      /dev_stage="fetal"
      1..1014
      /gene="DKFZp547F069"
      /gene="DKFZp547F069"
      /note="similarity to PEN11B (Homo sapiens)"
      /note_start=2
      /product="hypothetical protein"
      /protein_id="CAD38595.1"
      /db_xref="GI:21732472"
      /translation="ISLDEKEJFVLKDKPLISIKADIYAFISIPLSHVSLSOTS
      FRREYKASGSPVFOKRVROVDISSEGEPSRRGSGGGGITYFTTILSPSR
      FRKVFETIQALSTHDQPSVALADKNGAQTNPAGAPPSILOPPGPPDELSPP
      RRPDPKDKLALNCTPLP"
      968
      polyA_site
      /gene="DKFZp547F069"
      229 a 327 c 275 g 183 t
BASE COUNT
ORIGIN
Query Match      23.1%; Score 538; DB 9; Length 1014;
Best Local Similarity 100.0%; Pred. No. 2e-87;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

KEYWORDS
SOURCE      oligo capping; fls (full insert sequence).
            Homo sapiens ileal mucosa cDNA to mRNA, clone_11b:KAIA
            clone:KAIA00425.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      Kawakami,T., Noguchi,S., Itoh,T., Shigetani,K., Senba,T.,
            Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Tanigami,A.,
            Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hiroo,M.,
            Ohmori,Y., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T.,
            Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
            NEDO human cDNA sequencing project
            Unpublished
            2 (bases 1 to 3156)
TITLE
JOURNAL
REFERENCE
AUTHORS      Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
            Direct Submission
            Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure, Human Genome
            Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
            (E-mail:cdna@lems.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT
            NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing;
            Research Association for Biotechnology; cDNA library construction,
            5'- & 3'-end one pass sequencing; Department of Virology and Human
            Genome Center, Institute of Medical Science, University of Tokyo
            (partly supported by Science and Technology Agency).
FEATURES
            source
            1..3156
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="KAIA00425"
            /tissue_type="ileal mucosa"
            /clone_lib="KAIA"
            /note="Cloning vector: pME18SFL3"
BASE COUNT      588 a 1075 c 957 g 536 t
ORIGIN
Query Match      21.1%; Score 491.8; DB 9; Length 3156;
Best Local Similarity 61.9%; Pred. No. 4.1e-79;
Matches 1075; Conservative 0; Mismatches 372; Indels 290; Gaps 7;

```


Db	328	CAAGATGATGACCCGAAAC-----ACAGGACCCCATCTCTGGAAAGGTGGAGATCCGCTC	381
Qy	924	CCTGCCATTCAAACGGAGAGCTGGACCCCCGACGTCTTAGAGAGCATGGCATCTACTGGGCTG	983
Db	382	GCTGCCACACCTTGGAGGACATGACCCCGACGTCTGTGAGACAGCATGCATCTACTGGGCTG	441
Qy	984	CTTCAAGGACCGCGAGAGGCTGCTAGCGCGCTGCACAGTGAAGAGGAAACCAAGAAA	1043
Db	442	CTTCCGAGACCCCAACAACCTTGTGCAAGGACCTGTCTGCTCCGAGGAGGAAACCCGAGGAA	501
Qy	1044	GATGATATATATATCTGCTTTTGGATCGAAGAGACGGTATCCAGCTGTAGAGCAACGA	1103
Db	502	GATGATTTACTTCTCTCTCTGACCGGAAAGAAAGBTACCAGCCAGAGAGATGAGGA	561
Qy	1104	CTTGCCCTCCCGGAATGATGTTGACCCCCCGGAACGTGTGGATTCTCCATGTGAG	1163
Db	562	CTTGCCCCCGGAAAGSAGATAGACCTCTCCCGGAACGTGTGACTCCCGATGTGAA	621
Qy	1164	CCGTACGGGAAAGCGCGGACCGAGGGGAAAGTGCATGGAAGTCTGTAGCATACCAATGC	1223
Db	622	CCGGCAGCGCAAGCGCGGCGCCAGAAAGCAATCATGAGAGTGTCTAGCGTAGA-----	675
Qy	1224	CGGGGGTGTGTGCTCCCTCTGTACCCACCGGACGGGCTTGGAGATGGCCAGCACGA	1283
Db	676	---GGAGCGCGGCTCCCGGTGCTCTGCGGCGGGGCGCATTTGAGATGGCCACGACGGCA	732
Qy	1284	GAGATCCGTAGCGTCAGTGAGACCTTCACGGGTGTGTCTGTCCAGCCCTTAAAGACGCC	1343
Db	732	GAGGTCTCGGTTCATACAGCGGTGCTCTCTACAGSCCTTTCCACAGCCCACTACAGAGCC	792
Qy	1344	AAGAGATCGGCTCTTCTCTTTACCGGAGCGGGGGGCTGTGAGATAGGCTCGAGGCGG	1403
Db	792	CGGCG-----	797
Qy	1404	GGGCTCCCGACTTCAAAAAGCAGACGCTGCTTCGGGGGCCCAAGGGTGGGGGCGC	1463
Db	798	-----	797
Qy	1464	CGGGGAGCAGCCCCCGCCCGCAGTGCCCGGCTTCACACCCCTGCGCGGCCCCAGGCTC	1523
Db	798	-----TGACCCCTCACCCCTCACAAAGGGGCAGTCCCTCC-----	835
Qy	1524	CCCGCGCTCTCTGCGGGGACCCCTTTCAGCTGCTCTGACACAGCCCGGCGCACTGC	1583
Db	836	-----ACCCCCAAAGGGGACACTGTCTCACAGCGCAAAAGTAGACGCC	876
Qy	1584	CACCGGAGCCCCGGGGACAACAACACCCCGCAGCCCGGGGGGTGGCGTGGGGGAGGCCG	1643
Db	877	GGCTGGCAGCGCCCAACCCACGCGCCCCCTCTCACGCC-----AGCGTGGAGGGGTGCG	930
Qy	1644	CTGGAAGATCTCTTCAACTCCATCCATCCGCAACACTTCTTGGGCTCCCTCTGCTTTCACG	1703
Db	931	CTGGAGGGGGGGCTCAACTCCATCAAGAAGACTTCTTGGGCTCACCCGCTTCACAGC	990
Qy	1704	GGCGAAGATGACAGTCCCTAACCGCTGAGAGAGATGTCACGCTTGACGCGCAGATCTCCCC	1763
Db	991	CCGGAAGACTGAGTCCGACGCGGAGAGAGATGTCCAACCTGACACAGAGTGTCTCCCC	1050
Qy	1764	GGAGTGTGCAAAACGCTCTGTGTCTGGGAACTTCACTCTTGTGACAAAGAAACAAT	1823
Db	1051	AGAGTGTGGAAAGAGTCTGTGTTTGGAACTTCACTAGCTCTGGAGGAAGGAGGACGAT	1110
Qy	1824	ATTCTCTGCTAAAGAGCAAACTTTCAGCAGCATCAAGACAGACATGCTCATGCTT	1883
Db	1111	CTTCTGTGCTCAAAAGACAAACTTGTAGCTCTCACTACAGGCTGACATCTGTGACGCTT	1170
Qy	1884	TCTGTGATCCCGCAGCTGATGACAGTGTCTGTACAGACCAAGCTTCAAGGCGCAGTA	1943
Db	1171	CTGTGCTGATTCACAGTCTGACGCGACAGCGCATCTCCAAAGCAGCTTCCGGGCGGAGTA	1230
Qy	1944	CAAGGCGAGTGGCGGCGCTCGCTCTTCAAAAGCCCGCTCCAGTGTGACATCAG	2003

DB	1231	CAAGGCCAACGGGGGGGGGGCAGCGGTCTTCACGAAGCCGGTCAAGATTCCAGGTTGATATAC	1290
OY	2004	CTCCCTCTGAGAGGTCACAGACCCCTCCCCCGGAGCGGACGGCAGCGGAGAGTGTGTCATCTA	2063
DB	1291	CTACACGGAGAGGTT-----GGGAGGCGCCAGAAAGCAACGGCATCTTA	1332
OY	2064	CTCCGTCACCTTCACTCATCTGCGGGTCCAGCGCGTGGGTTCAAGCGAGTGGTGGAGAC	2123
DB	1333	CTCCGTCACCTTCACTCATCTGCGGGTCCAGCGCGTGGGTTCAAGCGAGTGGTGGAGAC	1392
OY	2124	CATCCAGGCGACACGCTCTGAGCAGCTCATGACAGCCCTCCGTCGAGCCCTGGCAGA	2180
DB	1393	CATCCAGGCGCGAGCTGTGAGACACAGACAGCCCGCTCGGGGCCAGCACTTGTCAGA	1449
RESULT 12			
AB014885			
LOCUS	AB014885	3933 bp	linear
DEFINITION	Halocynthia roretzi HrPOPK-1 mRNA, complete cds.		INV 06-FEB-1999
ACCESSION	AB014885		
VERSION	AB014885.1	GI:3172110	
KEYWORDS	HrPOPK-1.		
SOURCE	Halocynthia roretzi early embryonic egg tissue_lhb:fertilized egg		
ORGANISM	cdna to mRNA.		
REFERENCE	Enayayota; Metazoa; Chordata; Urochordata; Ascidiacea; Stolidobranchia; Pyuridae; Halocynthia.		
AUTHORS	1 (sites)		
TITLE	Sasakura, Y., Ogasawara, M. and Makabe, K. W.		
JOURNAL	Maternally localized RNA encoding a serine/threonine protein kinase		
MEDLINE	in the ascidian, Halocynthia roretzi		
REFERENCE	Mech. Dev. 76 (1-2), 161-163 (1998)		
AUTHORS	98440280		
TITLE	2 (bases 1 to 3933)		
JOURNAL	Sasakura, Y., Ogasawara, M. and Makabe, K. W.		
MEDLINE	Submitted (26-MAY-1998) Yasunori Sasakura, Kyoto University.		
REFERENCE	Department of Zoology, Graduate School of Science, Sakyo-ku, Kyoto,		
AUTHORS	Kuto 606-8502, Japan (E-mail: sasakura@ascidian.zool.kyoto-u.ac.jp,		
TITLE	Tel: +81-75-753-4095, Fax: +81-75-705-1113)		
JOURNAL	Location/Qualifiers		
FEATURES	1..3933		
SOURCE	/organism="Halocynthia roretzi"		
	/db_xref="taxon:7729"		
	/tissue_type="egg"		
	/dev_stage="early embryonic"		
	/tissue_lhb="fertilized egg"		
	1..3933		
	/gene="HrPOPK-1"		
	166..2262		
	/gene="HrPOPK-1"		
	/protein_start=1		
	/codon_start=1		
	/db_xref="GI:3172111"		
	/translation="MSNAPPOGCVGVPPYKLEKTLGKGGTGLVRLGVHMTGKRVAVKIK		
	VNRKLSESVINKEVERELAIIMKLEHHHTIGLHYVNRKXYIILILVSGELFDVIL		
	VOKRLLIPREARRFFROIISAVDYCHAHNHCCHRLKRENNLLDEKNNIKVADREMAST		
	QPESEFLLETSCGSHVACPEYIRGERYDGRADVSCGVLIFALLVGLPPDDNLLRQ		
	LLEKVRGKYVHIPIHFVPDQMLRGKVIDVAPDKRLSLQOVLQHPMWRPSSVYEGVIL		
	VTPEPVVVIDVCYHLPPEESYDPLVASMTSLGFCNKREKLKMLITEQVTEKVVYV		
	MLRRKKRYPFDDDDADSLPCBKDDADPRKRVDSYSLSSNGDMCVNPIIPORKMSAE		
	SICLTDSSSLPILSKRKSTEHOROSLTGSENSRLYONISPOKTAESKRINTGPPVRR		
	GTTCCSSNQPVQINTPPASPNPWRRLASLAKTTPGSGRPFRRHKQAPSSDEVENQMS		
	SSELSPKRCSEGNENMSSTSTSEHDELPALAIANKRLNSVKSCLVAHFLSIPVLTSHS		
	WVSFTRKRCDEGRSSGTSYTFVHQRSIKPFQVDIIQHSLSLDROENKGRPSQVGSFTPI		
	AFSLISGPIRKRYKVELIQLQWQMSIVAAANOONKTSGEFSAFSHEIIOOLCTVNNLHOD		
	NSTVAIINNPNHKSRYK"		
	3933		
	/gene="HrPOPK-1"		
	/note="19 A nucleotides"		
	/BASE COUNT	1138 a	803 c
	ORIGIN	751 g	1241 t

QY 379 AGAGTACGACCCCAAGAGGCGCCGAAAGTCTTCTCCGCGAGATGTGTCTGCGCTGACCTC 438
 DB 122 CGATTGACGCCGGAAGAGGCGCGGCAAGTCTTTCAGGCAAAATCATCTCGCCCTGGATTTC 181
 QY 439 TGGCAGACGCTACCTCTGCTGCGACAGAGACCTAAAGCCCGAAGCTCTTTGGATGAG 498
 DB 182 TGGCAGCTGCGATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 241
 QY 499 AAAAACAACATCCGCTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 558
 DB 242 AAGAAATACATTAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 301
 QY 559 CTGAGAGACGCTGCGCGGCTCCCGCATTTAGCTGTCTGAGATGAGATGAGATGAGATGAG 618
 DB 302 TTGAGAGACCTGCTGCGGCGACACATGACGCTGTCCAGAGATGATGAGATGAGATGAG 361
 QY 619 TATGATGAGCGCGCGGCGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 678
 DB 362 TACGATGAGCGCGGCGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 421
 QY 679 GGGGCTGCGCTTGTGATGAGACGACATCTCCGAGCTGTGAGAGATGAGATGAGATGAGATGAG 738
 DB 422 GGTGCGCTGCGCTTGTGATGAGACGACATCTCCGAGCTGTGAGAGATGAGATGAGATGAG 481
 QY 739 GTCTTCCACATGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 798
 DB 482 GTCTTCCACATGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 541
 QY 799 GAAGTGAAGCCGCAAAAAGGCTGAGTGAAGATGAGATGAGATGAGATGAGATGAGATGAG 858
 DB 542 GAGGTCAATCGGAGCGCGGCTGAGTGAAGATGAGATGAGATGAGATGAGATGAGATGAG 601
 QY 859 GGGGCGAAGACGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 918
 DB 602 GCGGCGGCGCAAGG---GAGCTGAGCTGAGCTGCCAATGATGAGATGAGATGAGATGAG 658
 QY 919 CGGAGCTGCGCATCAAGGAGAGCTGAGACCCGAGCTGCTGAGAGATGAGATGAGATGAG 978
 DB 659 CAGGTATTCACAGGCGCGGCTGAGTCCGATGCTGAGAGATGAGATGAGATGAGATGAG 718
 QY 979 GGGCTGCTGAGGCGGAGAGAGCTGAGTCCGAGCTGAGAGATGAGAGATGAGAGATGAG 1038
 DB 719 GCGCTGCTGAGGAGAGAGAGATGAGTCCGAGCTGAGAGATGAGATGAGATGAGATGAG 778
 QY 1039 GAAAAGATGATATATATATCTGCTTTGGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1098
 DB 779 GAGAGGTTATATATATCTGCTTTGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 838
 QY 1099 CAGGACCTGCTCCCGGAA 1118
 DB 839 GATGATGAGATGCGCAAA 858
 RESULT 15
 AC008974 38000 bp DNA linear PRI 26-JAN-2001
 LOCUS AC008974 Homo sapiens chromosome 19 clone LTNLF-116A8, complete sequence.
 DEFINITION AC008974
 ACCESSION AC008974
 VERSION AC008974.7 GI:12545283
 KEYWORDS HTG.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 38000)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 38000)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 3 (bases 1 to 38000)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Submitted (26-JAN-2001) DOE Joint Genome Institute, 2800 Mitchell
 Drive, Walnut Creek, CA 94598, USA
 COMMENT On Jan 26, 2001 this sequence version replaced gi:7711382.
 Draft Sequence Produced by DOE Joint Genome Institute
 www.jgi.doe.gov
 www.shgc.stanford.edu
 Quality: Phrap Quality >=40 99.8% of Sequence;
 Estimated Total Number of Errors is 0.2.
 FEATURES
 source
 1. 38000
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19"
 /clone="LTNLF-116A8"
 BASE COUNT 8785 a 9978 c 10552 g 8685 t
 ORIGIN
 Query Match 15.9%; Score 371; DB 9; Length 38000;
 Best Local Similarity 100.0%; Pred. No. 2.2e-57;
 Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1347 GAGTCCGGCTTTCTTCTTTCACCGAGCGGCGGCTGAGATGAGCTCGAGCGGGG 1406
 DB 33273 GAGTCCGGCTTTCTTCTTTCACCGAGCGGCGGCTGAGATGAGCTCGAGCGGGG 33332
 QY 1407 CTCCCGGCTTCCAAAAGCGAGAGCTGCTTCTGCGGCGGCGGCGGCGGCGGCGG 1466
 DB 33333 CTCCCGGCTTCCAAAAGCGAGAGCTGCTTCTGCGGCGGCGGCGGCGGCGGCGG 33392
 QY 1467 GAGAGAGCG 1526
 DB 33393 GAGAGAGCG 33452
 QY 1527 GCGCTCTCTGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1586
 DB 33453 GCGCTCTCTGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 33512
 QY 1587 GCGGAGCGCGCGGAGACACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1646
 DB 33513 GCGGAGCGCGCGGAGACACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 33572
 QY 1647 GAGAGAGCTCTCAACTCCATCCGCAAGAGCTTCTGCGGCTGCGCTTCAAGCGCG 1706
 DB 33573 GAGAGAGCTCTCAACTCCATCCGCAAGAGCTTCTGCGGCTGCGCTTCAAGCGCG 33632
 QY 1707 CAAGATGACAG 1717
 DB 33633 CAAGATGACAG 33643

Search completed: April 21, 2003, 23:05:39
 Job time: 4079.82 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2003, 12:29:03 ; Search time 19 Seconds
(without alignments)
1204.791 Million cell updates/sec

Title: US-10-003-690-2

Perfect score: 4108
Sequence: 1 MSGGAKGGGGSPAYHLPHP.....PRKGPXKRLATNGTLP 778

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCBUS.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2799	68.1	668	4	US-09-930-181-2
2	2490.5	60.6	603	4	US-09-930-181-17
3	2385.5	58.1	585	4	US-09-930-181-4
4	827	20.1	745	4	US-09-523-849-36
5	795	19.4	729	2	US-08-677-398-2
6	795	19.4	729	4	US-09-523-849-33
7	788	19.2	733	4	US-09-523-849-32
8	784	18.3	776	4	US-09-523-849-34
9	753.5	18.1	633	4	US-08-557-006C-43
10	713.5	17.4	345	3	US-09-101-146-1
11	715	17.4	552	4	US-08-557-006C-40
12	700.5	17.1	257	3	US-09-101-146-6
13	695	16.9	604	4	US-09-523-849-35
14	691	16.8	257	2	US-07-857-224B-25
15	570	14.1	149	4	US-09-930-181-18
16	543.5	13.2	222	2	US-07-857-224B-26
17	522.5	12.7	436	4	US-09-734-673-2
18	521	12.7	436	4	US-09-523-849-2
19	521	12.7	1050	4	US-09-428-711A-16
20	518.5	12.6	353	3	US-08-688-988-31
21	512	12.5	290	4	US-09-734-673-4
22	512	12.5	334	4	US-09-523-849-31
23	510	12.4	1037	4	US-09-428-711A-21
24	509.5	12.4	260	2	US-07-857-224B-27
25	495	12.0	339	3	US-08-688-988-33
26	493	12.0	351	3	US-08-688-988-28
27	491	12.0	354	3	US-08-688-988-29

28	490.5	11.9	260	2	US-07-857-224B-28	Sequence 28, Appl
29	482.5	11.7	370	2	US-08-878-989-19	Sequence 19, Appl
30	482.5	11.7	370	4	US-09-272-796-19	Sequence 19, Appl
31	482.5	11.7	370	4	US-09-457-040B-31	Sequence 31, Appl
32	482.5	11.7	1051	4	US-09-428-711A-14	Sequence 14, Appl
33	482	11.7	556	4	US-09-800-960-4	Sequence 4, Appl
34	477.5	11.6	264	2	US-07-857-224B-18	Sequence 18, Appl
35	477.5	11.6	359	3	US-08-688-988-32	Sequence 32, Appl
36	477	11.6	363	3	US-08-688-988-30	Sequence 30, Appl
37	475.5	11.6	565	4	US-09-800-960-2	Sequence 2, Appl
38	474.5	11.6	264	2	US-07-857-224B-19	Sequence 19, Appl
39	469.5	11.4	295	1	US-07-951-715A-23	Sequence 23, Appl
40	469.5	11.4	295	2	US-08-459-448A-23	Sequence 23, Appl
41	469.5	11.4	295	3	US-08-459-595A-23	Sequence 23, Appl
42	469.5	11.4	295	3	US-08-459-504B-23	Sequence 23, Appl
43	469.5	11.4	295	3	US-08-459-444-23	Sequence 23, Appl
44	469.5	11.4	295	4	US-09-547-422-23	Sequence 23, Appl
45	467.5	11.4	339	3	US-08-688-988-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-930-181-2
Sequence 2, Application US/09930181
Patent No. 6455292
GENERAL INFORMATION:
APPLICANT: Origene Technologies
TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
FILE REFERENCE: 160 101 VI
CURRENT APPLICATION NUMBER: US/09/930,181
CURRENT FILING DATE: 2001-08-16
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 668
TYPE: PRT
ORGANISM: Homo sapiens
US-09-930-181-2

Query Match 68.1%; Score 2799; DB 4; Length 668;
Best Local Similarity 75.1%; Pred. No. 8.9e-171;
Matches 348; Conservative 54; Mismatches 48; Indels 80; Gaps 8;

QY	1	MSGGAKGGGGSPAYHLPHPHPPOHAQYGPRLKTKGOTGLVKGHCYTGOKV	60
DB	1	MTSTGADGA-----QHAQYGPRLKTKGOTGLVKGHCYTGOKV	45
QY	61	AIKIVNREKLSSEVLMKVEREIALTKLIEHPVILKLDHYENKKYLYLVLEHVSGLLEFD	120
DB	46	AIKIVNREKLSSEVLMKVEREIALTKLIEHPVILKLDHYENKKYLYLVLEHVSGLLEFD	105
QY	121	YLVRKGLTPREKARKFRQVLSALDFCHSYSTICRDLKPEWLLDDEKNNIRADFGASL	180
DB	106	YLVRKGLTPREKARKFRQVLSALDFCHSYSTICRDLKPEWLLDDEKNNIRADFGASL	165
QY	181	QVDSLEETSCGSPHYACPEYIKGEKDGRRADMMSCGVILFALLVGLPDDDDLRQL	240
DB	166	QVDSLEETSCGSPHYACPEYIKGEKDGRRADMMSCGVILFALLVGLPDDDDLRQL	225
QY	241	EKKVKGVEHMPHFLPPDCQSLRGMIEVEPERKLSLEQIOKHPWYLGKHEPDPLEBAP	300
DB	226	EKKVKGVEHMPHFLPPDCQSLRGMIEVEPERKLSLEQIOKHPWYLGKHEPDPLEBAP	284
QY	301	GRVYARSLPSNGELDDVLESASLGCFRPRELHRELSSEENQEMITLYLLDRER	360
DB	285	GRVYARSLPSNGELDDVLESASLGCFRPRELHRELSSEENQEMITLYLLDRER	343
QY	361	YPCSEDDLPDPRNDVDPKRRKVDSPMLSRHGRPKRPSMEVLSITDAGGSGSPVPTARA	420
DB	344	YPCSEDDLPDPRNDVDPKRRKVDSPMLSRHGRPKRPSMEVLSITDAGGSGSPVPTARA	400

```

QY 421 LEMAHORSRSVSGASTGLSSPLSSPRPVFSFSPERGAGDEARGGGSPRTKQTLP 480
DB 401 LEMAHGORSRSISGASSGLSTSPSSP----- 429
QY 481 RGRGCGAGCEPPPSARSPLPGPPSPSSGGTPLHPLHTPRASPTGTPPTPPSP 540
DB 430 -----VTPHPSRGSPLPTPKG-----TPVHTPKESAGTPNTPPPSP 468
QY 541 GGGVGAAMRSRLNSIRNSFLGSPRHRKMKQVPTAEEMSSLTPESSPELAKRSWFGNTI 600
DB 469 --SVGVPMFARLNSIKNSFLGSPRHRKMKQVPTPEEMSSLTPESSPELAKRSWFGNTI 526
QY 601 SLDKESQIFLYLKDPLSSIKADIVHAFSTPSSHSVLSQTSFRAEKAGSGSPVFOKP 660
DB 527 SLEKEQIFLYLKDPLSSIKADIVHAFSTPSSHSVLSQTSFRAEKAGSGSPVFOKP 586
QY 661 VREOVDISSEGEPEPPRRRSGGGGIYSVTFTLISGSPRRKRVETIQAOILSTHOP 720
DB 587 VKFOVDITTEGEQAQK-----NGIYSVTFTLISGSPRRKRVETIQAOILSTHOP 640
QY 721 SVQALDEKN 730
DB 641 AAQHLSDTTN 650

```

RESULT 2

```

US-09-930-181-17
; Sequence 17, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 V1
; CURRENT APPLICATION NUMBER: US/09/930,181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 603
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-930-181-17

```

```

Query Match 60.6%; Score 2490.5; DB 4; Length 603;
Best Local Similarity 72.8%; Pred. No. 3.6e-151;
Matches 488; Conservative 52; Mismatches 45; Indels 85; Gaps 9;

```

```

QY 87 LLEHPRVTLKHVDYENKKYLYLVLEHVSGLGELPDYLYKKGRLTPKEAKRFRQIVSALDF 146
DB 1 LLEHPRVTLKHVDYENKKYLYLVLEHVSGLGELPDYLYKKGRLTPKEAKRFRQIVSALDF 60
QY 147 CHSYISCHRDLPENLLDEKNNIRIADFGMASLQVDSILETSCGSPHYACPEVIKGEK 206
DB 61 CHSHSICHDRDKENLLDEKNNIRIADFGMASLQVDSILETSCGSPHYACPEVIKGEK 120
QY 207 YGRRADAMSCGYILFALLVGAALPPDDDNIRQLLEKVKRGVFMHPHPIPPDCQSLRGMI 266
DB 121 YGRRADAMSCGYILFALLVGAALPPDDDNIRQLLEKVKRGVFMHPHPIPPDCQSLRGMI 180
QY 267 EYEPERLLEJOIKHPWLYGKHHPDPCLEPAAGRVAMRSLSNGLDIDVDVLESASLT 326
DB 181 EYDAAARLLEHIOKHIWYIGKNEPEP-RQPIR-RKVQIRSLPSLEDIDVDVLESASLT 238
QY 327 GCFRDERLARELSEENOEKMIYLLIDRKERYPSCEDODLPPRNDVDPKRKYDSPM 386
DB 239 GCFRDERLARELSEENOEKMIYLLIDRKERYPSCEDODLPPRNDVDPKRKYDSPM 298
QY 387 LSRHGKRRPERKSMYVLSITDAGGGSPVPTRRALEMAHSORSRSVSGASTGLSSPLS 446
DB 299 LNRHGKRRPERKSMYVLSITD--GGSPVPAARALEMAHORSRSISGASSGLSTSPLS 355
QY 447 SFRSPVFSFSPERGAGDEARGGGSPRTKQTLPSSRGRGGAGCEPPPSARSPTLP 506
DB 447 SFRSPVFSFSPERGAGDEARGGGSPRTKQTLPSSRGRGGAGCEPPPSARSPTLP 506

```

```

DB 356 SPR-----VTPHPSRGSPLPTPK 374
QY 507 GSPRSAGTPLHPLHTPRASPTGTPPTPPSPGGVGAAMRSRLNSIRNSFLGSPRF 566
DB 375 G-----TPVHTPKESAGTPNTPPPSP--SVGVPMFARLNSIKNSFLGSPRF 421
QY 567 HRRKMQVPTAEEMSSLTPESSPELAKRSWFGNTISLDEKQIFLYLKDPLSSIKADIVH 626
DB 422 HRRKMQVPTAEEMSSLTPESSPELAKRSWFGNTISLDEKQIFLYLKDPLSSIKADIVH 481
QY 627 AFLTSLSHSVLSQTSFRAEKAGSGSPVFOKPYRVOVDISSEGEPEPPRRRSGGGG 686
DB 482 AFLTSLSHSVLSQTSFRAEKAGSGSPVFOKPYRVOVDITTEGEQAQK-----NG 535
QY 687 IYSVTFTLISGSPRRKRVETIQAOILSTHOPSVQALDEKNAGOTRPAAGAPRSLQ 746
DB 536 IYSVTFTLISGSPRRKRVETIQAOILSTHOPPAQHLS-----EP 577
QY 747 PPRGRPELS 756
DB 578 PP--PAPGLS 585

```

RESULT 3

```

US-09-930-181-4
; Sequence 4, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 V1
; CURRENT APPLICATION NUMBER: US/09/930,181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-930-181-4

```

```

Query Match 58.1%; Score 2385.5; DB 4; Length 585;
Best Local Similarity 74.4%; Pred. No. 1.7e-144;
Matches 464; Conservative 51; Mismatches 44; Indels 65; Gaps 7;

```

```

QY 107 YVLEHVSGLGELPDYLYKKGRLTPKEAKRFRQIVSALDFCHSYISCHRDLPENLLDE 166
DB 9 YVLEHVSGLGELPDYLYKKGRLTPKEAKRFRQIVSALDFCHSYISCHRDLPENLLDE 68
QY 167 KNNIRIADFGMASLQVDSILETSCGSPHYACPEVIKGEKYDGRADAMSCGYILFALLV 226
DB 69 KNNIRIADFGMASLQVDSILETSCGSPHYACPEVIKGEKYDGRADAMSCGYILFALLV 128
QY 227 GALPDDDNIRQLLEKVKRGVFMHPHPIPPDCQSLRGMIYEPERLLEJOIKHPWYL 286
DB 129 GALPDDDNIRQLLEKVKRGVFMHPHPIPPDCQSLRGMIYEPERLLEJOIKHPWYL 188
QY 287 GSKHHPDPCLEPAAGRVAMRSLSNGLDIDVDVLESASLGCFRDERLARELSEENQ 346
DB 189 GSKHHPDPCLEPAAGRVAMRSLSNGLDIDVDVLESASLGCFRDERLARELSEENQ 246
QY 347 EKMITYLLIDRKERYPSCEDODLPPRNDVDPKRKYDSPMLSHRGKRRPERKSMYVLSIT 406
DB 247 EKMITYLLIDRKERYPSCEDODLPPRNDVDPKRKYDSPMLSHRGKRRPERKSMYVLSIT 306
QY 407 DAGGGSPVPTRRALEMAHSORSRSVSGASTGLSSPLSSPRPVFSFSPERGAGDEAR 466
DB 307 D--GGSPVPAARALEMAHORSRSISGASSGLSTSPSSP----- 346
QY 467 GGGSPRTKQTLPSSRGRGGAGCEPPPSARSPTLPGRGSPRSAGGTPLHPLHTPRA 526
DB 347 -----VTPHPSRGSPLPTPKG-----TPVHTPKESAGTPNTPPPSP 371

```



```

QY 527 SPTGTPPTPPSPGGVGAAMRSRLNSINSLGSPFRHRRKQVPTAEMSSLTIPS 586
DB 372 SPATPTPTPPSP--SVGGVPMRARRLNSINSLGSPFRHRRKQVPTAEMSSLTIPS 429
QY 587 SPPLAKSMWGNFSLKREDFIYLKDKPLSTIKADIVHAFLSPISHSVLSOTSPFA 646
DB 430 SPPLAKSMWGNFSLKREDFIYLKDKPLSTIKADIVHAFLSPISHSVLSOTSPFA 489
QY 647 EYKAGGPPSVFQKRVPRQVVDISSEGEPPSPRRDGGSGGCGIYVTFLLISGPRRRKRV 706
DB 490 EYKATGGPAVFQKRVKQVVDITTEGEBAQK-----NGIYVTFLLISGPRRRKRV 543
QY 707 EYKATGGPAVFQKRVKQVVDITTEGEBAQK-----NGIYVTFLLISGPRRRKRV 543
DB 544 EYKATGGPAVFQKRVKQVVDITTEGEBAQK-----NGIYVTFLLISGPRRRKRV 543

RESULT 4
US-09-523-849-36
; Sequence 36, Application US/09523849
; Patent No. 6458561
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Molteni, Angela
; APPLICANT: Magnaghi, Paola
; APPLICANT: Bosotti, Roberta
; APPLICANT: Scaccheri, Emanuela
; APPLICANT: Isacchi, Antonella
; APPLICANT: Hodgson, Dave
; TITLE OF INVENTION: HUMAN NIM1 KINASE
; FILE REFERENCE: PC-0009 US
; CURRENT APPLICATION NUMBER: US/09/523, 849
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PERL Program
; SEQ ID NO 36
; LENGTH: 745
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank Accession No. 6458561 g1749794
US-09-523-849-36

Query Match          20.1%; Score 827; DB 4; Length 745;
Best Local Similarity 30.9%; Pred. No. 4,5e-45;
Matches 245; Conservative 114; Mismatches 259; Indels 176; Gaps 26;

QY 30 YGPPYRLKTLGKGTGLVLCVHCITGQKVAIKIVNREKLSVLMKVEREIAIKLIE 89
DB 16 HIGYRLKTLGKGNFPAKVLARHILGKEVAVKIIDTQLNSSLQKLFREVRIMKVLN 75
QY 90 HPVHLKHDVYENKYYLVLEHVSQGLFPIYVKKGLTPREARKFPQIYSAIDFCS 149
DB 76 HPVHLKHDVYENKYYLVLEHVSQGLFPIYVKKGLTPREARKFPQIYSAIDFCS 135
QY 150 YSICHRDLKPELILDKENNRIRIADFGMASLQVDSLETSCGSPHYACPEVYIKGEKYD 209
DB 136 KFIYHRDLKPELILDKENNRIRIADFGMASLQVDSLETSCGSPHYACPEVYIKGEKYD 195
QY 210 RRDWMSGCVLFLVLLYALFPDDNLFQLLLEKRYGVHMPHETPPDCQSLIRGMEVE 269
DB 196 RRDWMSGCVLFLVLLYALFPDDNLFQLLLEKRYGVHMPHETPPDCQSLIRGMEVE 255
QY 270 PEKRLSLEQLOKHFWYVGGKHEPD---PCLPAPGRVAMRSLPNSGELDPVLGSMASL 326
DB 256 PEKRLSLEQLOKHFWYVGGKHEPD---PCLPAPGRVAMRSLPNSGELDPVLGSMASL 302
QY 327 GCFDRRLRLNELNSEENQKMYLLLLDRKERYPCEDDLPPIRNDV-----DPRTELAVSM 376
DB 303 GCFDRRLRLNELNSEENQKMYLLLLDRKERYPCEDDLPPIRNDV-----DPRTELAVSM 359
QY 377 -----PPRRKRVDS-----PMLSRHGK-----RRPE-----RKSM----- 400

```

```

DB 360 HKVQSVANPKNORRFSDOAGPALPTNSYSKKTQSNNAENKRRPEEDRESGRKASSTAKV 419
QY 401 -----EVLISITDAGGGGSPVPTPRALEMA--OHSORSVSG--A 436
DB 420 PASPLPRLERKKTTPPTSTNSVLSITNRSNRPULLEPASIGQASIQNGKSLTPGSR 479
QY 437 STLSSPLSPR-----SPVFSFSPGAGDEARGGSPSTKQTLPSRGP----- 483
DB 480 STASASAAVSAARPRQKMSASVHNNKASGLPPTSENCVPRSTAPQRPVAPSAH 539
QY 484 ---RGAGAGEQ---PPPSARSTPLRPGSPRSQGGTGPLSLPHTPAASPTGTPTTP 537
DB 540 NISSGAPDRPTNFPKGVSSSTFHAQLRQVROOQNLPT-----GVV-P 583
QY 538 PSPGGVGAAMRSRLNSINSLGSPFRHRRKQVPTAEE-MSLTP----- 584
DB 584 ASPSGHSGR--KGASSTISKF--TSKTVARNLNEPSKRVETLRPHVYGGGNDKEX 639
QY 585 ---ESSPELAKRW--EGNTSLDKEDQIFLYLKDPLSTIKADIVHAFLSPISHSVLS 639
DB 640 EEPREAPRSLRFTWSMKTSSMEPNEMREIRRYLDANSQSELHEKYMILL--CMHGTP 697
QY 640 SOTSFRAEYKASGSPVQKRVPRQVVDISSEGEPPSPRRDGGSGGCGIYVTFLLISGPR 699
DB 698 GHEDF-----VQMEMEVCK-----LPRLSLNG-----VFKRISGTS 729
QY 700 RRRFRVETIOQL 713
DB 730 MAFKNASKTANEL 743

```

```

RESULT 5
US-08-677-298-2
; Sequence 2, Application US/08677298
; Patent No. 5863729
; GENERAL INFORMATION:
; APPLICANT: Pivnicka-Worms, Helen
; TITLE OF INVENTION: DNA SEQUENCES ENCODING HUMAN TCAR-1
; TITLE OF INVENTION: KINASE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: CO
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk.
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/677, 298
; FILING DATE: 09-JUL-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Caruthers, Jennie M.
; REGISTRATION NUMBER: 34,464
; REFERENCE/DOCKET NUMBER: 9-96
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 729 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-677-298-2

```

Query Match 19.4%; Score 795; DB 2; Length 729;

Best Local Similarity 31.9%; Pred. No. 4.8e-43; Matches 239; Conservative 109; Mismatches 262; Indels 140; Gaps 28;

QY	30	YVGPRLEETLKEGGOTGLVKLGHVCHTIGOVAKIYNREKLSSTYLMKVEEITLILIE	89
Db	52	HIGNRLKLTICKGNFAKKYKLAHILITGEBVALITIKDOLNPSLCKREVAIMKTLN	111
QY	90	HPHVLLHDVYENKKYLYLVEHVSGGELFEDYLVKKGRILPKPEARKKRPJOIVSALDCHS	149
Db	112	HPNIVKLEFEVIEKTLTYLIMEVASGEFVDYVANGRMKEKERSKRFROIVSAVOYCHQ	171
QY	150	YSICHRDLKPEUMLDEKKNIRIADFGMASLOYDLSLETCGSPHYACBPAVIGKERYDG	209
Db	172	KRIYVRDILKAEMLLDADANNIKIADGFCSENEFYGGKIDTFCGSPPAABELEFGKKYDG	231
QY	210	KRADWMSGVILFALLVGMALPDDOOLRLLEKVKGVGHVEMHPFIPDDCSILRGMIEVE	269
Db	232	PEVDWMSLGVILTYLTVSGSLPDDGOMLKELEKREVLNGKTRIPIYMSDTCENLKLRFVLN	291
QY	270	PEKRLSEIOIKHPWYLGKHNED--PCLPAPGRHVARSISPNGEL--DPDVLISM	323
Db	292	PIKRGTLRQIMKDMRINMG-HEDEKLPPVEP-----ELDISQKRIDIM	335
QY	324	ASLGCFFRDERLHRELRESEENO--KITY-----YLLDKREKRYPCEDOD-----	368
Db	336	VGMG-----YSOEEIOESTLSKKMYDEITATYLLGKRKSSLELDADSSSSNLS	383
QY	369	---LPPRRDYPD-----PPKRVDSPLSHHGKRPRPKSMSEVLSTIDAGGCGSP-----VP	416
Db	384	LAKVPPSSDLNASTGOSPHHKVORSVSSOCKOR-----YSDHAPALPVSAYP	433
QY	417	TRRALEMAHOSR-----SRSVGASGTGLSSPFLSPRSPVFESFSPGAGDEARGCGSP	471
Db	434	KRSQSTFADGLKEKDQISSRKSSSGSAVGKG--LAPASMLGNASPNKADIDE-----	485
QY	472	TSKYUTLDSKRGCGAGAQPPPPKARSTPLPQPPSPSSGGTPLHSLHTPPRASFTGT	531
Db	486	RKKSSTVSPSNATASGM-----TRRNTYV--CSERTTAOR--HSVJONGKENST-I	531
QY	532	PQTPPPSPGGGVGGAAMRSRL-----NSIRNSFTLSPRRHRRKMQVPTAEKMSLTPES	586
Db	532	PDQRPVPVASTHSSISNAATPDRIRFPGRJASSTFHHQPR-ERRTATYINGPPASPISLSHEA	590
QY	587	SPELAKRS-----WFGNFIS-LDKERQILFVLKD-----KPLS-----SIK-----AD	623
Db	591	TPLSQTRSGNSTPLSKLSKILTRSNVNSAEDQDENKKAAPSLRLFTWSMKTITSSMDPGD	650
QY	624	IVHAPLSTPSLSHVSLSQTSFPAETKASGSPVYQKPRVRFQVNDISSSEGPSPDRDSCG	683
Db	651	MMREIRKLYLDANNCYEGORE-REFLFCVAGHDGHAENMLVOMEMEVCCK-----LPRSLNG	703
QY	684	GGGIYSTFTLISGSPSRKFRVETIOALD	713
Db	704	-----VPRKISGTSLAFKNITASKIANEL	727

RESULT 6
US-09-523-849-33

```

; SOFTWARE: PERL Program
; SEQ ID NO 33
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: GenBank Accession No. 6458561.93089349
; US-09-523-849-33

```

Query Match	Similarity	19.48	Score	795	DB	43	Length	729
Best Local	Similarity	31.99	Pred. No.	4.8e-43				
Matches	239	Conservative	109	Mismatches	262	Indels	140	Gaps
Qy	30	YVGPRLREKTLGKQGTGLVYKLVGHVCHITGQKAKIIVNREKLSSESVLKMKVEREIKILKLE	89					
Db	52	HIGNVRLFKLTIGKGNFAKVKFLARHILHGREVAIKITDKQTQNPSTLQKLFREVRIMILN	111					
Qy	90	HPHYLKLHDVYENKKYLLVLEHVSGLLEFYLYKGRHLPRKPKRFFRQIVASLDLFGHS	194					
Db	112	HPNIVYKLEFEVETEKTYLLIMEYASGGVEPDYLAHGMKEKERSKFRQIVASVVOYCHQ	171					
Qy	150	YSICRDLKPENLILDEKNNIRIDFGAASLOVDSLSLETSCGSPHACPFVINGEYTDG	209					
Db	172	KRIYHRDLKANNLLDDMDNMIKIDFGESNFEYTGKLDITFCGSPPIAAPELFOGKATYDG	233					
Qy	210	RRADMSCGVTLFALLVGLALPEDDDNIRQLLEKYKGVFHMHPHIPPDCOSLNGMIEVE	265					
Db	232	PEVWVMSGLVILYTLVSSLSLPEFDQNLKELEBRLRGKRYRIFPMYSIDCEMLTKRFLVN	291					
Qy	270	PEKRLSLEIOIKHWMYLGKKEHP---PCLEPAGRGVAMKSLPSNEL---PDVLESM	323					
Db	292	PIKGTLEQLIKKDMINNG-HEDELKPYEP-----ELDISQKRIDIM	335					
Qy	324	ASLSCFDRERHARELSEENOE---KMYT-----YLIDRKERYPCEDOD-----	368					
Db	336	VGMG-----YSQEIQISLSMKKDEITATYLLILGRKSSELDASDSSSNLS	383					
Qy	369	---LPPRADDV---PPKRVDSPLMSRHKRRPERKSMEVLSTDAGGGSP---VP	416					
Db	384	LAKVAPSSDLNNSTGQSPHHNVQASVSSQOKOR-----YSHAGPAIPSYVAYP	433					
Qy	417	TRRLAEMAHQSOR-----SREVSASAGLSSPLSPSPFPFSPPBGADGDERGGSP	471					
Db	434	KRSQSTIRADGLKEDGJSSKRSSSSAYGKG---IAASPLGLGASNPKNADPE-----	485					
Qy	472	TSKTQTLTBSRPGGAGGEODPPPSABSTPLPBPBGSPRSSGGTPLHSPHTAPASTGT	531					
Db	486	RKKSIVYSSWTAGGM-----TRNTYV---GSERTADR--HVIYONGKENST-I	531					
Qy	532	PGTTPPSPGCGVGGAAMKRTL-----NSIRNSLGSPPRHRRKMQVPTLEMSLTPES	566					
Db	532	PDQRTPAVASTHSSAATPDRIIRPRTJASTHGHQPR-ERRATYVNGPPASPSLSHEA	590					
Qy	587	SPELAIRS-----WFGNFI-SLDKEEOIFVLKD-----KPLS-----SLK-----AD	623					
Db	591	TPLSQTRSROSTNLFSLKTLSTLSTRVNVSAEQDKNKEAPRSLRTFTWSKATITSSMDPCD	630					
Qy	624	IVAHFLSLPSLSHSVLSQTSFRAEYKASGGSYFQKPFVRQYVDISSSEGPSPRRDGG	683					
Db	651	MMREIRKIVLDANNQDYEQRE-RELFECVHGHAENLVOWEMEYCK-----LPRLSLNG	703					
Qy	684	GGGIYVTFLLISGSPRRFRKRVETIQAOQL	713					
Db	704	-----VFRKRISGTSLAFKATSKINLEL	727					

RESULT 7
US-09-523-849-32
; Sequence 32, Application US/09523849
; Patent No. 6458561
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga


```

Db 513 FSASEGAGLGGGATGTTGCTSSFPVLASPLGS-----QSAPVLOSQGLGATVPL 566
OY 589 -----ELAKRSWFGNFISLDEKQIPLVLDKPLSSIKADI 624
Db 567 PVSHQEGRRASDTSLTGGLKAFRQGLRKNARTKGFGLGK----- 606
OY 625 VHAFLSPISLASHVLSQTSFPAEKASGSPVQKPVPRQVDISSECEPSPRRDGS 684
Db 607 -----TKGLARQV-COSSING-----SRGSMSTFHTPA-----PSSGLG--- 640
OY 685 GGATSVFTLLSGSPRRKRVETI--QAOLSTHDDPSVQALADEKNGAOTRPAAGAPR 742
Db 641 -----CTASSREGRLSEVLHMQRLDLOHSAVS--SDYQAPQLSPV---PY 685
OY 743 SLOPPG 749
Db 686 VLFPCDG 692

```

RESULT 9

```

US-08-557-006C-43
; Sequence 43, Application US/08557006C
; Patent No. 6258547
; GENERAL INFORMATION:
; APPLICANT: Bertl, Rajindar K.
; APPLICANT: Carling, David
; APPLICANT: Forde, Robert A.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
; FILE REFERENCE: NGAF/PHM37588/US
; CURRENT APPLICATION NUMBER: US/08/557,006C
; CURRENT FILING DATE: 1996-03-06
; PRIOR APPLICATION NUMBER: PCT/GB94/01093
; PRIOR FILING DATE: 1994-05-20
; PRIOR APPLICATION NUMBER: GB 9310489.1
; PRIOR FILING DATE: 1993-05-21
; PRIOR APPLICATION NUMBER: GB 9318010.7
; PRIOR FILING DATE: 1993-08-31
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 43
; LENGTH: 633
; TYPE: PRT
; ORGANISM: Yeast
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(633)
; OTHER INFORMATION: Yeast SNF1 polypeptide
US-08-557-006C-43

```

```

Query Match 18.3%; Score 753.5; DB 4; Length 633;
Best local Similarity 32.8%; Pred. No. 1.8e-40;
Matches 166; Conservative 103; Mismatches 144; Indels 93; Gaps 14;

```

```

OY 14 AYHLPRHPHPPROHAQ-----YVGPYREKLTGKOTGLVKGVCNI 55
Db 17 SHHHHHHHHHHHHGGGNSLTNNPKSSLDGAHIGNYIVTKTLEGGSPGVKLAUHTT 76
OY 56 TGOQVAIKIYNREKLSSEVLM-KVERETAILKLEHPVLKLDHYENKKYUYLEHVS 114
Db 77 TGOQVAIKIYNREKLSSEVLM-KVERETAILKLEHPVLKLDHYENKKYUYLEHVS 114
OY 115 GGELFDLVKGRITPREARKFRQIVSALDFCHSYSTICHDLPENLLDEKNNIRIAD 174
Db 136 GNEFLDIYVRDMSSEGEARFFQOIIISAVEYCHRHKIYHDLKPEMLLDLHNVKIAD 195
OY 175 FGMAISLOVGSILETSCGSPHYACPEVIKYGKRGADMSGCIYIEALVGALEPDD 234
Db 196 FGISNITDGNFCKTSCGSPHYACPEVIKYGKRGADMSGCIYIEALVGALEPDD 234
OY 235 NLQMLEKVKRGVFMHFTIPPCDSSLGMIYEPEKRLSLEQIKHPWYLGKHE-- 291
Db 256 SIYVLFKRNISNGVYTLKFLSPGAAGLIKRLILYPLNRISIHIMODWFRVYDLPETLL 315

```

```

OY 292 -PDPCLEPAPGRRYAMR-----SLPSNGELDPVLSMASLGGFRDRERLRLRSEEE 344
Db 316 PPD--LKPHDEEENENDSKKDGSSPDNDEIDMIVNLSSTMGV-EKDEIYESLESSED 372
OY 345 -----NOKEMTYILLDDKREKRPSCEDDILPRRNV-----DPP-----RK 380
Db 373 TPAPNEIRDAIVMLTKENKSLI---KDMKANKSVSDLEDTFLSGSPPTFOQSKSHOKS 427
OY 381 RVDSFPLSRHGRKRPERKSMSEVLSTDAGGSGSPVPRFALE---MAQHSQSRSVSGA 436
Db 428 QVHEFAKQARNR-----ASAKITQORTYHQSFPMDQYKHEEDSTVSL 470
OY 437 STGL-----SSSPLSPNSPV 452
Db 471 PTLSPQIHRANMLAOGSPAASKISPL 496

```

RESULT 10

```

US-09-101-146-1
; Sequence 1, Application US/09101146
; Patent No. 6124125
; GENERAL INFORMATION:
; APPLICANT: Dartmouth College, St. Vincents Institute of
; APPLICANT: Medical Research, Kemp et al.
; TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 66 E. Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/101,146
; FILING DATE: October 7, 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PN7450
; FILING DATE: 8 JAN 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: DC-0050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (856) 810-1515
; TELEFAX: (856) 810-1454
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 345
; TYPE: Amino acid
; TOPOLOGY: linear
US-09-101-146-1

```

```

Query Match 17.4%; Score 715.5; DB 3; Length 345;
Best local Similarity 40.7%; Pred. No. 2.3e-38;
Matches 146; Conservative 65; Mismatches 105; Indels 43; Gaps 5;

```

```

OY 31 VGPYREKLTGKOTGLVKGVCNITGQVAKIVNREKLSSEVLMYEREITATIKLEH 90
Db 12 IGHVILGDTLGVSTFGVKGKHELTHKVAVKTLNRQKRLADVAGKIRREIQMLKLRH 71
OY 91 PHTLKLHDYENKKYUYLEHVS GGELFDLVKGRITPREARKFRQIVSALDFCHSY 150
Db 72 PHIKIKLYVISTSDIFNMVAYVSGELFDYICNGRLDEKSRRLDQQLISGVYDCHRH 131
OY 151 STCHRLKPEMLLLDEKNNIRIADFGMAISLOVGSILETSCGSPHYACPEVIKGEKYGDR 210

```

```

Db      132  MVAHRDLKPNVLLDAMNNAKIDAFGLSNMMSDGEFLTSCGSPVYAPAEVYISGLYAGP 191
OY      211  RADMWSGVLLFALLVAGALPPDDNLRQLLEKVGVEFHMFFIPDCOSLIRGIEVEP 270
Db      192  EVDIMSSGVLLALCTLPFDDHVPFLFKKIDGIFTYTOYINPVSISLKHMLQYDP 251
OY      271  EKRSLSEQIOKHPWYLGKHEPDPCLPAPGRVAMRSLPSNGELDPVLESMAISLGCGR 330
Db      252  MKRATIKDIREHEWF---KODLPKYLFPPE-----DPSYSTMI----- 286
OY      331  DRERLRRLRELRSEENOKMI-----YIILLDRKERYPSCEDODL---PP 371
Db      287  DDEALKEVCEKFCSESEVLSCLYNNHODPLAVAYHLIDNRIRIMNAKDPYLATSP 345

```

RESULT 11
US-08-557-006C-40

```

; Sequence 40, Application US/08557006C
; Patent No. 6238347
; GENERAL INFORMATION:
; APPLICANT: Berl, Rajindar K.
; APPLICANT: Carling, David
; APPLICANT: Forde, Robert A.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
; FILE REFERENCE: NGAP/PM37588/05T
; CURRENT APPLICATION NUMBER: US/08/557, 006C
; PRIOR FILING DATE: 1996-03-06
; PRIOR APPLICATION NUMBER: PCT/GB94/01093
; PRIOR FILING DATE: 1994-05-20
; PRIOR APPLICATION NUMBER: GB 9310489.1
; PRIOR FILING DATE: 1993-05-21
; PRIOR APPLICATION NUMBER: GB 9318010.7
; PRIOR FILING DATE: 1993-08-31
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 40
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Rat
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)-(1747)
; OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -
; OTHER INFORMATION: fragment begins at nucleotide 24 and ends with
; OTHER INFORMATION: nucleotide 1765
US-08-557-006C-40

```

Query Match 17.4%; Score 715; DB 4; Length 552;

Best Local Similarity 40.7%; Pred. No. 4.3e-36; Matches 150; Conservative 71; Mismatches 134; Indels 14; Gaps 5;

```

OY      31  VGYRLKTKTGKGTGVLKGVHCITGOKVAIKIYVREKL-SESYLAKVEREITALIKLIE 89
Db      13  IGHVILGDTLGVGTFGKVGKGEHOLTGKAVKILNRKISLDVVGKIKREIQMLKLEFR 72
OY      90  HPHVLIKLDVYENKKYLLVLEHVSSELFDYLVKKGRTPREARKFROIYSALDFCHS 149
Db      73  HPHIKLYQVISTPDPFVMEYVSGSELDYICKHGRVEVEARLRLOQLISANDYCHR 132
OY      150  YSICHRDLKPNLILDEKNNIRIDAFGNASIQVDSLETSCGSPHYACPEYIKGEKIDG 209
Db      133  HMYVHRDLKPNVLLDAMNNAKIDAFGLSNMMSDGEFLTSCGSPVYAPAEVYISGLYAGP 192
OY      210  RADMWSGVLLFALLVAGALPPDDNLRQLLEKVGVEFHMFFIPDCOSLIRGIEVEP 269
Db      193  PVDIMSSGVLLALCTLPFDDHVPFLFKKIRGVFTYIPYINRSIATILMMLQYDP 252
OY      270  PEKRLSLEQIOKHPWYLGKHEPDPCLPAPGRVAMRSLPSNGELDPVLESMAISLGCGR 329
Db      253  PLKRAITIKDIREHEWF---KODLPKYLFPPE-----PSYANVYIDAEVAKVCEKFCGR 302
OY      330  RDRELRRLRELRSEENOKMIYIILLDRKERYPSCEDODL---PPRNDVDPKRYVDSPPM 386

```

```

Db      303  TESEVNNISYSGPQODLAVAYHLIDNRIRIMNAQASEFYIASSPPTGSGFMDMAMHIPP 362
OY      387  LSRHGRKRP 395
Db      363  LKHPHERMP 371

```

RESULT 12
US-09-101-146-6

```

; Sequence 6, Application US/09101146
; Patent No. 6124125
; GENERAL INFORMATION:
; APPLICANT: Dartmouth College, St. Vincents Institute of
; APPLICANT: Medical Research, Kemp et al.
; TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 66 E. Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM PC
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/101,146
; FILING DATE: October 7, 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PN7450
; FILING DATE: 8 JAN 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: DC-0050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (856) 810-1515
; TELEFAX: (856) 810-1454
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257
; TYPE: Amino acid
; TOPOLOGY: Linear
US-09-101-146-6

```

Query Match 17.1%; Score 700.5; DB 3; Length 257;

Best Local Similarity 50.6%; Pred. No. 1.5e-37; Matches 129; Conservative 52; Mismatches 73; Indels 1; Gaps 1;

```

OY      32  GPYRLKTKTGKGTGVLKGVHCITGOKVAIKIYVREKL-SESYLAKVEREITALIKLIEH 90
Db      1  GHVILGDTLGVGTFGKVGKGEHOLTGKAVKILNRKISLDVVGKIKREIQMLKLEFR 60
OY      91  HPHVLIKLDVYENKKYLLVLEHVSSELFDYLVKKGRTPREARKFROIYSALDFCHS 150
Db      61  PHIKLYQVISTPDPFVMEYVSGSELDYICKHGRVEVEARLRLOQLISANDYCHR 120
OY      151  SICHRDLKPNLILDEKNNIRIDAFGNASIQVDSLETSCGSPHYACPEYIKGEKIDG 210
Db      121  MVAHRDLKPNVLLDAMNNAKIDAFGLSNMMSDGEFLTSCGSPVYAPAEVYISGLYAGP 180
OY      211  RADMWSGVLLFALLVAGALPPDDNLRQLLEKVGVEFHMFFIPDCOSLIRGIEVEP 270
Db      181  EVDIMSSGVLLALCTLPFDDHVPFLFKKIDGIFTYTOYINPVSISLKHMLQYDP 240
OY      271  EKRSLSEQIOKHPWYLGKHEPDPCLPAPGRVAMRSLPSNGELDPVLESMAISLGCGR 329
Db      241  MKRATIKDIREHEWF 255

```



```
Db 120 KIVARDLKPENLLDDEHINVAIADPGLSNIMTDGNEFLKTSQGSFNYAAPEVLSKLIYAGP 179
QY 211 RADMMSCGVLLFALLVGLPEPDDNLRLQLEKVKRGVFMHPHFIPDCOSLRGMIEVEP 270
Db 180 EVDVWSCGVLLYVLCRLPEPDDSEIPVLFKNISNGVYTLPLFSLPGAAGLIKRMILYVP 239
QY 271 EKRLSEIQIOKHPPY 285
Db 240 LNRISIHIMDDWF 254
```

RESULT 15

```
US-09-930-181-18
; Sequence 18, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 160 101 V1
; CURRENT APPLICATION NUMBER: US/09/930.181
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 149
; TYPE: PT
; ORGANISM: Homo sapiens
US-09-930-181-18
```

Query Match

14.1%; Score 579; DB 4; Length 149;

Best Local Similarity 80.7%; Pred. No. 4.2e-30;

Matches 117; Conservative 13; Mismatches 7; Indels 8; Gaps 2;

```
QY 579 MSSTPSSPELAKRSWGFNSIDKEQIFLYLKDPLSSIKADIVHAFSLIPSLSHSV 638
Db 1 MSNLTPESSPELAKRSWGFNSIDKEQIFVYIKDKPLSSIKADIVHAFSLIPSLSHSV 60
QY 639 LSQTSFRAEYKASGSPVFOKPVRFQVDISSSEGEPPSRDGGSGGCIYSVFTLLISGP 698
Db 61 ISQTSFRAEYKATGPAVFOKPVRFQVDITYTEGGEAQKE-----NGIYSVFTLLISGP 114
QY 699 SRRKRVETIQAOILSTHD--QPS 721
Db 115 SRRKRVETIQAOILSTHDPLRPS 139
```

Search completed: April 16, 2003, 12:33:27
Job time : 24 secs

GenCore version 5.1.4_P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2003, 12:28:03 ; Search time 98 Seconds
(without alignments)
1635.762 Million cell updates/sec

Title: US-10-003-690-2
Perfect score: 4108
Sequence: 1 MSSGAKGEGGSPAYHLPHP.....PRRGPPKKLLATNGTPLP 778

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPREMBL.21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4108	100.0	778	4	08TDC2	08TDC2 homo sapien
2	3956.5	96.3	794	4	08TDC3	08TDC3 homo sapien
3	3758	91.5	715	4	096J14	096J14 homo sapien
4	2490.5	60.6	603	4	060843	060843 homo sapien
5	2466	47.5	851	5	096AV4	096AV4 homo sapien
6	1952	42.5	698	5	061298	061298 caenorhabdi
7	1814	44.2	914	5	019469	019469 drosophila
8	1744	42.5	701	5	095T82	095T82 drosophila
9	1228	29.9	833	5	08SSX5	08SSX5 dictyosteli
10	874	21.3	1371	4	09Y2K2	09Y2K2 homo sapien
11	832	20.3	691	4	096RG0	096RG0 homo sapien
12	828	20.2	745	4	015524	015524 homo sapien
13	827	20.1	755	4	096HB3	096HB3 homo sapien
14	827	20.1	688	4	09BYD8	09BYD8 homo sapien
15	826	20.0	689	4	096JG7	096JG7 homo sapien
16	823	20.0	689	4	096JG7	096JG7 homo sapien

17	823	20.0	752	4	096L34	096L34 homo sapien
18	822	20.0	722	11	008679	008679 rattus norv
19	814	19.8	1192	5	09TW45	09TW45 caenorhabdi
20	814	19.8	1193	5	017346	017346 xenopus lae
21	812.5	19.7	1096	5	017368	017368 caenorhabdi
22	809	19.6	1096	5	017368	017368 caenorhabdi
23	806.5	19.6	729	4	09P0L2	09P0L2 homo sapien
24	796	19.4	729	4	09P0L2	09P0L2 homo sapien
25	795.5	19.4	535	10	038MD2	038MD2 homo sapien
26	795	19.4	729	4	060219	060219 homo sapien
27	793	19.3	401	8	098SC8	098SC8 guillardi
28	792.5	19.3	504	10	093113	093113 guillardi
29	792	19.3	744	4	096RG1	096RG1 homo sapien
30	789	19.2	729	11	09J0K4	09J0K4 mus musculu
31	789	19.2	926	4	09H0K1	09H0K1 mus musculu
32	788.5	19.2	744	11	09J0K5	09J0K5 mus musculu
33	788	19.2	753	4	08W8X3	08W8X3 mus musculu
34	788	19.2	793	11	008678	008678 rattus norv
35	786	19.1	514	10	09X225	09X225 mus musculu
36	784	19.1	795	11	08VH45	08VH45 mus musculu
37	783.5	19.1	514	10	08VH45	08VH45 mus musculu
38	781	19.0	797	11	08VH45	08VH45 mus musculu
39	780.5	19.0	798	13	091A88	091A88 rattus norv
40	777	18.9	505	10	09ZRJ1	09ZRJ1 oryza sativ
41	776	18.9	938	5	09Y8V8	09Y8V8 drosophila
42	772	18.8	1075	5	095U75	095U75 drosophila
43	770.5	18.8	752	4	09UN34	09UN34 homo sapien
44	767	18.7	1398	5	077268	077268 drosophila
45	767	18.7	1398	5	09W532	09W532 drosophila

ALIGNMENTS

RESULT 1	PRELIMINARY:	PRT:	778 AA.
08TDC2			
1	08TDC2		
AC	01-JUN-2002 (TREMURel. 21, Created)		
DT	01-JUN-2002 (TREMURel. 21, Last sequence update)		
DT	01-JUN-2002 (TREMURel. 21, Last annotation update)		
DE	Protein kinase-like protein.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID-9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	She X.Y., Yu L., Guo J.H.;		
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AF479827; AAL87698.1;		
KW	kinase.		
SC	SEQUENCE 778 AA: 85086 MW: 8D1818DAE54398BB CRC64;		

Query Match	Best Local Similarity	100.0%	Score 4108:	DB 4:	Length 778:
Matches	778:	Conservative	0:	Mismatches	0:
QY	1	MSSGAKGEGGSPAYHLPHPPOHAYVGYRLKTLGKGTGLVKGVCITGQKV	60		
DB	1	MSSGAKGEGGSPAYHLPHPPOHAYVGYRLKTLGKGTGLVKGVCITGQKV	60		
QY	61	AIKIVNREKLSVLMKVERETAILKLHPHYLKHVDYENKKYLYLLEHVSSEGLPD	120		
DB	61	AIKIVNREKLSVLMKVERETAILKLHPHYLKHVDYENKKYLYLLEHVSSEGLPD	120		
QY	121	YLVKGRLLPKPKKRFKRFQIVSLDFCHSYICHRDLKPEMLLDEKNNIRIADFQMASL	180		
DB	121	YLVKGRLLPKPKKRFKRFQIVSLDFCHSYICHRDLKPEMLLDEKNNIRIADFQMASL	180		
QY	181	QVDSLTLETSCGSPHYACPEVIRGKYYDGRADMSGCVILFALLVGALEPDDDNIROLL	240		
DB	181	QVDSLTLETSCGSPHYACPEVIRGKYYDGRADMSGCVILFALLVGALEPDDDNIROLL	240		

```

QY 241 EKVRGVFHMHPFIPPCOSILRGMEVEPEKRLSLDIOKHWPYLGKHEPDCLEPAP 300
DB 241 EKVRGVFHMHPFIPPCOSILRGMEVEPEKRLSLDIOKHWPYLGKHEPDCLEPAP 300
QY 301 GRRVAMSLNSNGELDDVLESMAISLCCFDRERLHRELSEENOEKMIYLLDRKER 360
DB 301 GRRVAMSLNSNGELDDVLESMAISLCCFDRERLHRELSEENOEKMIYLLDRKER 360
QY 361 YPSCEDODLPPRDVDPFRKRVDSFMLSRGKRPERKSMENVLSTPDAGGSGSVPRRA 420
DB 361 YPSCEDODLPPRDVDPFRKRVDSFMLSRGKRPERKSMENVLSTPDAGGSGSVPRRA 420
QY 421 LEMAHOSRSRVSASTGSLSSPSPSPVFSFSPGAGDEARCGSPSTKOTLPS 480
DB 421 LEMAHOSRSRVSASTGSLSSPSPSPVFSFSPGAGDEARCGSPSTKOTLPS 480
QY 481 RSPRGAGAGEOPPPPARSTPLPGPPSPSSGCTPLHSPHTPRASPTGTPPPSP 540
DB 481 RSPRGAGAGEOPPPPARSTPLPGPPSPSSGCTPLHSPHTPRASPTGTPPPSP 540
QY 541 GGGVGAAMRSRLNSIRNSFLGSPRHRKMOVPTAEEMSLTPESSPELAKSMFGNFI 600
DB 541 GGGVGAAMRSRLNSIRNSFLGSPRHRKMOVPTAEEMSLTPESSPELAKSMFGNFI 600
QY 601 SLDRKEQIFLVKDKPLSLIKADIVHAFSTPSLSHVSLSQTSFRAEYASGSPVFORP 660
DB 601 SLDRKEQIFLVKDKPLSLIKADIVHAFSTPSLSHVSLSQTSFRAEYASGSPVFORP 660
QY 661 VRFQVDISSSEPPSPRDGSGGGIYSVFTLLISGPRRKRVRVETIOALLSTHDP 720
DB 661 VRFQVDISSSEPPSPRDGSGGGIYSVFTLLISGPRRKRVRVETIOALLSTHDP 720
QY 721 SVQALADEKNAQTRPAPAPRSLQPPGRDPELSSSPRRGPPDKKLATNGTPLP 778
DB 721 SVQALADEKNAQTRPAPAPRSLQPPGRDPELSSSPRRGPPDKKLATNGTPLP 778

```

RESULT 2

```

ID 08TDC3 PRELIMINARY; PRT; 794 AA.
AC 08TDC3;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Putative serine/threonine protein kinase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RA She X.Y., Guo J.H., Yu L.;
RL Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF479826; AAL8797.1;
KW Kinase; Serine/threonine-protein kinase.
SQ SEQUENCE 794 AA; 86753 MW; 5DD395B0E61AEF77 CRC64;

```

Query Match 96.3%; Score 3956.5; DB 4; Length 794;

Best Local Similarity 98.1%; Pred. No. 7.8e-254;

Matches 756; Conservative 1; Mismatches 9; Indels 5; Gaps 1;

```

QY 8 GGGGSPAYHLPRHPHPORHAQYVGPYRLKTLGKQGLVNLGVHCITGQVVAIKIYVR 67
DB 29 GGGGSPAYHLPRHPHPORHAQYVGPYRLKTLGKQGLVNLGVHCITGQVVAIKIYVR 67
QY 68 EKLSSEYVMKVEREALIKLIEHPVLYLHDYENKKYLYLVLEHVSGLPFDYVYKGR 127
DB 84 EKLSSEYVMKVEREALIKLIEHPVLYLHDYENKKYLYLVLEHVSGLPFDYVYKGR 127
QY 128 LTPKARKEFRQIVSALDFCHSYSTCHDLPENLLDEKNINRIADGMAISLOVGDSTL 187
DB 128 LTPKARKEFRQIVSALDFCHSYSTCHDLPENLLDEKNINRIADGMAISLOVGDSTL 187

```

```

DB 144 LTPKARKEFRQIVSALDFCHSYSTCHDLPENLLDEKNINRIADGMAISLOVGDSTL 203
QY 188 ETSQSPYACPEVYIKGKTYGRADWMSGVILFALLVGALEPDDDLROLLEKVRGV 247
DB 204 ETSQSPYACPEVYIKGKTYGRADWMSGVILFALLVGALEPDDDLROLLEKVRGV 263
QY 248 FHMHPFIPPCOSILRGMEVEPEKRLSLDIOKHWPYLGKHEPDCLEPAPGRVAMR 307
DB 264 FHMHPFIPPCOSILRGMEVEPEKRLSLDIOKHWPYLGKHEPDCLEPAPGRVAMR 323
QY 308 SLPSNGELDDVLESMAISLCCFDRERLHRELSEENOEKMIYLLDRKERPSCEDQ 367
DB 324 SLPSNGELDDVLESMAISLCCFDRERLHRELSEENOEKMIYLLDRKERPSCEDQ 383
QY 368 DLPPRDVDPFRKRVDSFMLSRGKRPERKSMENVLSTPDAGGSGSVPRRALEMAQHS 427
DB 384 DLPPRDVDPFRKRVDSFMLSRGKRPERKSMENVLSTPDAGGSGSVPRRALEMAQHS 443
QY 428 QRSRSVSGASTGLSSPSPSPVFSFSPGAGDEARCGSPSTKOTLPSRGRGG 487
DB 444 QRSRSVSGASTGLSSPSPSPVFSFSPGAGDEARCGSPSTKOTLPSRGRGG 503
QY 488 ACEQPPPARSTPLPGPPSPSSGCTPLHSPHTPRASPTGTPPPSPGCGVGA 547
DB 504 ACEQPPPARSTPLPGPPSPSSGCTPLHSPHTPRASPTGTPPPSPGCGVGA 563
QY 548 AMRSRLNSIRNSFLGSPRHRKMOVPTAEEMSLTPESSPELAKSMFGNFI 607
DB 564 AMRSRLNSIRNSFLGSPRHRKMOVPTAEEMSLTPESSPELAKSMFGNFI 623
QY 608 IFLVLDKRLSLIKADIVHAFSTPSLSHVSLSQTSFRAEYASGSPVFORP 667
DB 624 IFLVLDKRLSLIKADIVHAFSTPSLSHVSLSQTSFRAEYASGSPVFORP 683
QY 668 SSSSEPPSPRDGSGGGIYSVFTLLISGPRRKRVRVETIOALLSTHDP 727
DB 684 SSSSEPPSPRDGSGGGIYSVFTLLISGPRRKRVRVETIOALLSTHDP 743
QY 728 EKNQATRPAPAPRSLQPPGRDPELSSSPRRGPPDKKLATNGTPLP 778
DB 744 EKNQATRPAPAPRSLQPPGRDPELSSSPRRGPPDKKLATNGTPLP 794

```

RESULT 3

```

ID 096JL4 PRELIMINARY; PRT; 715 AA.
AC 096JL4;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE KIAA1811 protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RA Nagase T., Nakayama M., Nakajima D., Kitano R., Ohara O.;
RL The complete of the coding sequences of 100 new cDNA clones from brain which code
for large proteins in vitro.
DNA Res. 8:85-95(2001).
EMBL; AB058714; BAB47440.1;
InterPro: IPR000719; Euk. pk. kinase.
InterPro: IPR002290; Ser. thr. kinase.
pfam: PF00069; pk. kinase. 1.
ProDom: PD000001; Euk. kinase. 1.
PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.

```

FT NON_TER 1 1
 SQ SEQUENCE 715 AA; 78499 MW; B90FEE115C418A5 CRC64;
 Query Match
 Best Local Similarity 100.0%; Pred. No. 9,6e-241;
 Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 IYNRKESVSLAKVERIALIKLIEHPVLYKLHVEYENKYYLYLVEHVSGLDFLYL 123
 DB 1 IYNRKESVSLAKVERIALIKLIEHPVLYKLHVEYENKYYLYLVEHVSGLDFLYL 60
 QY 124 KKGRLTPKARKFFROIVSALDFCHSYISICHRIKPEMLLDEKNNIRIADFGMAISLOVG 183
 DB 61 KKGRLTPKARKFFROIVSALDFCHSYISICHRIKPEMLLDEKNNIRIADFGMAISLOVG 120
 QY 184 DSLLETSGSPHYACPEVYIKGEKYDGRADMGSCVILFALLVGLPPDDNLRLQLEKY 243
 DB 121 DSLLETSGSPHYACPEVYIKGEKYDGRADMGSCVILFALLVGLPPDDNLRLQLEKY 180
 QY 244 KGVFHMHPFIIPDCQSLRGMIYEVEPERKLSLEQIQKHPTVLGKHEPDCLEPAAGRR 303
 DB 181 KGVFHMHPFIIPDCQSLRGMIYEVEPERKLSLEQIQKHPTVLGKHEPDCLEPAAGRR 240
 QY 304 VAMRSLPSNGELDPVLESMAISLGCPRDERLHRELSEENOEKMIYLLDRKERYPS 363
 DB 241 VAMRSLPSNGELDPVLESMAISLGCPRDERLHRELSEENOEKMIYLLDRKERYPS 300
 QY 364 CEDODLPNRNDVPPRKRVDSPLSRHGRKRPERSMEVLSITDAGGSPVPTRALEM 423
 DB 301 CEDODLPNRNDVPPRKRVDSPLSRHGRKRPERSMEVLSITDAGGSPVPTRALEM 360
 QY 424 AQSQRSRVSGASTGLSSPLSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 483
 DB 361 AQSQRSRVSGASTGLSSPLSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 420
 QY 484 RGGAGAGQPPPSARSTPLPGPPSPSSSGTPLHSPPLTPRASPTGTPPTPPSPSGG 543
 DB 421 RGGAGAGQPPPSARSTPLPGPPSPSSSGTPLHSPPLTPRASPTGTPPTPPSPSGG 480
 QY 544 VGGAMRSRLNLSINSLGSPRRHRRKQVYTAEMSLTPESPSPELAKSMWENFTSLD 603
 DB 481 VGGAMRSRLNLSINSLGSPRRHRRKQVYTAEMSLTPESPSPELAKSMWENFTSLD 540
 QY 604 KEEQIFLYLKKPLSLIKADIYHAFSLPSLSHVSLSQTSFRAEYKASGGSPVQKPVRF 663
 DB 541 KEEQIFLYLKKPLSLIKADIYHAFSLPSLSHVSLSQTSFRAEYKASGGSPVQKPVRF 600
 QY 664 QVDISSSEGPSPRRDGGGGGGLYSVTFTLISGSRFRKVVETIOAQLSTHDQPSVQ 723
 DB 601 QVDISSSEGPSPRRDGGGGGGLYSVTFTLISGSRFRKVVETIOAQLSTHDQPSVQ 660
 QY 724 ALADKNGAQTAPAGAPPSLQPPGRDPPELSSSPRRKQPKKLLATNGTPLP 778
 DB 661 ALADKNGAQTAPAGAPPSLQPPGRDPPELSSSPRRKQPKKLLATNGTPLP 715

RESULT 4
 060843 PRELIMINARY: PRT: 603 AA.

AC 060843;
 DT 01-AUG-1998 (TEMBLrel. 07, last sequence update)
 DT 01-AUG-1998 (TEMBLrel. 20, last annotation update)
 DE Putative serine/threonine protein kinase (Fragment).
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RX NCBI_TaxID=9606;
 RN SEQUENCE FROM N.A.
 RP TISSUE=BRNIN;
 RC MEDLINE=21064499; PubMed=11124703;
 RA Stanchi F., Bertocco E., Topo S., Dioguardi R., Simonati B.,

RA Cannata N., Zimbello R., Lanfranchi G., Valle G.; high similarity to
 "Characterization of 16 novel human genes showing high similarity to
 yeast sequences";
 RT Yeast 18:69-80(2001).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AJ006701; CA007196.1; --
 DR HSSP: 063450; 1A06.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; pkinase.1.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Euk_pkinase.1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
 DR PROSITE: PS00108; PROTEIN_KINASE_SF.1.
 KW Kinase; Serine/threonine-protein kinase.

FT NON_TER 1
 SQ SEQUENCE 603 AA; 67401 MW; B02C5D678F8E96 CRC64;
 Query Match
 Best Local Similarity 72.8%; Pred. No. 7.1e-157;
 Matches 488; Conservative 52; Mismatches 45; Indels 85; Gaps 9;

QY 87 LIEHPVLYKLHVEYENKYYLYLVEHVSGLDFLYLKKGRLPKARKFFROIVSALDF 146
 DB 1 LIEHPVLYKLHVEYENKYYLYLVEHVSGLDFLYLKKGRLPKARKFFROIVSALDF 60
 QY 147 CHSYISICHRIKPEMLLDEKNNIRIADFGMAISLOVGSILETSGSPHYACPEVYIKGEK 206
 DB 61 CHSYISICHRIKPEMLLDEKNNIRIADFGMAISLOVGSILETSGSPHYACPEVYIKGEK 120
 QY 207 YGRRADMGSCVILFALLVGLPPDDNLRLQLEKYRGVYHMFHTIPDCQSLRMI 266
 DB 121 YGRRADMGSCVILFALLVGLPPDDNLRLQLEKYRGVYHMFHTIPDCQSLRMI 180
 QY 267 EYVEPERKLSLEQIQKHPTVLGKHEPDCLEPAAGRRVAMRSLPSNGELDPVLESMAISL 326
 DB 181 EYVEPERKLSLEQIQKHPTVLGKHEPDCLEPAAGRRVAMRSLPSNGELDPVLESMAISL 238
 QY 327 GCFRDERLHRELSEENOEKMIYLLDRKERYPSCEDODLPNRNDVPPRKRVDSM 386
 DB 239 GCFRDERLHRELSEENOEKMIYLLDRKERYPSCEDODLPNRNDVPPRKRVDSM 298
 QY 387 LSRHGRKRPERSMEVLSITDAGGSPVPTRALEMAQHSQSRVSGASTGLSSPLS 446
 DB 299 LSRHGRKRPERSMEVLSITDAGGSPVPTRALEMAQHSQSRVSGASTGLSSPLS 355
 QY 447 SPRSPVSP 506
 DB 356 SPRSPVSP 374
 QY 507 GSPRSSGTPPLHSPPLTPRASPTGTPPTPPSPSGGVGGAAMRSRLNLSINSLGSPRF 566
 DB 375 GSPRSSGTPPLHSPPLTPRASPTGTPPTPPSPSGGVGGAAMRSRLNLSINSLGSPRF 421
 QY 567 HRRKQVYTAEMSLTPESPSPELAKSMWENFTSLDKEQIFLYLKKPLSLIKADIYH 626
 DB 422 HRRKQVYTAEMSLTPESPSPELAKSMWENFTSLDKEQIFLYLKKPLSLIKADIYH 481
 QY 627 AFLSPSLSHVLSQTSFRAEYKASGGSPVQKPVROVYDSSSGSPSPRRDGGGGG 686
 DB 482 AFLSPSLSHVLSQTSFRAEYKATGPAVYQKPVROVYDSSSGSPSPRRDGGGGG 535
 QY 687 IYSVTFTLISGSRFRKVVETIOAQLSTHDQPSVQALADKNGAQTAPAGAPPSLQPP 746
 DB 536 IYSVTFTLISGSRFRKVVETIOAQLSTHDQPSVQALADKNGAQTAPAGAPPSLQPP 577
 QY 747 PPGRPPELS 756
 DB 578 PPGRPPELS 585

RESULT 5

```

096ANV4 ID 096ANV4 PRELIMINARY; PRT; 473 AA.
AC 096ANV4;
DT 01-DEC-2001 (TREMBLER, 19, Created)
DT 01-DEC-2001 (TREMBLER, 19, Last sequence update)
DT 01-JUN-2002 (TREMBLER, 21, Last annotation update)
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
RX NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC016681; AAHL6681.1;
DR InterPro: IPR000449; UBA_domain.
KW Hypothetical protein.
SQ
SEQUENCE 473 AA; 50750 MW; D27DAD437DA96AA9 CRC64;

Query Match
Best Local Similarity 60.0%; Score 2466; DB 4; Length 473;
Matches 472; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

0Y 306 MRSLSNGELDPVLESMASLCGCFDRERHRELSSEENDEKMTYLLDRKERYPSC 365
1 MRSLSNGELDPVLESMASLCGCFDRERHRELSSEENDEKMTYLLDRKERYPSC 60
DB 366 DODLPKRDVPPPKRVDSPMLSRHRRPKRMEVLSITDAGGSGPYTRALEMAQ 425
61 DODLPKRDVPPPKRVDSPMLSRHRRPKRMEVLSITDAGGSGPYTRALEMAQ 120
0Y 426 HSORSRSVSGASTGLSSPSLSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 485
121 HSORSRSVSGASTGLSSPSLSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 180
DB 486 GGAEPGPPPPARSTPPLPGPGSPSSSGTPLHSPLTPASPTGTPGTPPPSPGGG 545
181 GGAEPGPPPPARSTPPLPGPGSPSSSGTPLHSPLTPASPTGTPGTPPPSPGGG 240
0Y 546 GAAMRSRLNSIRNSFLSPFRHRRKQVPTAEEMSLTPRESSSELAKRSMFGNFI 605
241 GAAMRSRLNSIRNSFLSPFRHRRKQVPTAEEMSLTPRESSSELAKRSMFGNFI 300
DB 606 EQLFVLKPKPLSSIKADIVHAFSLPSLSHVSLSQTSFRAEYKASGGSVOKPVR 665
301 EQLFVLKPKPLSSIKADIVHAFSLPSLSHVSLSQTSFRAEYKASGGSVOKPVR 360
0Y 666 DISSSGPPSPRRDGGGGGYSVTFTLISGSPRRFRKAVETIQOILSTHDQPSVAL 725
361 DISSSGPPSPRRDGGGGGYSVTFTLISGSPRRFRKAVETIQOILSTHDQPSVAL 420
DB 726 ADEKGAQTRPAGAPRSLQPPGAPDELSPPRRGPKDKKLATNGTTPP 778
421 ADEKGAQTRPAGAPRSLQPPGAPDELSPPRRGPKDKKLATNGTTPP 473

RESULT 6
09VIVA 09VIVA PRELIMINARY; PRT; 851 AA.
AC 09VIVA;
DT 01-MAY-2000 (TREMBLER, 13, Created)
DT 01-MAY-2000 (TREMBLER, 13, Last sequence update)
DT 01-MAR-2002 (TREMBLER, 20, Last annotation update)
DE CG6114 protein.
GN CG6114.
OS Drosophila melanogaster (fruit fly).
NC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
NC Pterygota; Neoptera; Endopterygota; Diptera; Brachyoptera; Muscomorpha;
NC Phyllophaga; Drosophilidae; Drosophila.
RX NCBI_TaxID=7227;

RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazer J.R., Chapple M., Pfeiffer B.D.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burtis K.C., Busam D.A., Butler H., Cavalli L.P., Center A.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jallat M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Ralner T.K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden Klamas I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Weissman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1. SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AE003529; AAF9569.1;
DR HSSP: Q63450; 1A06.
DR FLYBASE: FB90036544; CG6114.
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR002290; Ser_Thr_Pkinase.
DR InterPro: IPR001245; Tyr_Pkinase.
DR Pfam: PF00069; Pkinase.1.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Euk_Pkinase.1.
DR SMART: SM00220; S_TKc.1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
DR PROSITE: PS00108; PROTEIN_KINASE_DOM.1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST.1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ
SEQUENCE 851 AA; 94152 MW; 9E5456E8F5D47E0 CRC64;

Query Match
Best Local Similarity 54.9%; Score 1952; DB 5; Length 851;
Matches 421; Conservative 96; Mismatches 138; Indels 112; Gaps 24;

0Y 26 QHQQYVGRLEKTLKGGTGLVKLGCHITGQVAKIKVNEKLSSEYLMKVEREINAIL 85
10 ENQGFVGRLEKTLKGGTGLVKLGCHITGQVAKIKVNEKLSSEYLMKVEREINAIL 69
DB 86 KLIEHPVLTLDVYENKRYLYLVLEHVSGLLEFDYLVKKGRLPKARKKRFQVSLAD 145
70 KLIDHNVLTLDVYENKRYLYLVLEHVSGLLEFDYLVKKGRLPKARKKRFQVSLAD 129
0Y 146 FCHSYSTCHRLKPNLLDEKNNRIADFGMASLQVDSLETSGSPHACPEVIRGE 205
130 FCHSHSTCHRLKPNLLDEKNNRIADFGMASLQVDSLETSGSPHACPEVIRGE 189

```


206 KYDGRADWMSGVILFALLVGLALPDDNLRKLEKRGVEMHPRFPDCCSLRGM 265
 190 KYDGRADWMSGVILFALLVGLALPDDNLRKLEKRGVEMHPRFPDCCSLRGM 249
 266 IEVEPERKLSLEOIKHPMT-LOGKHEPD---PCLPAPGRVAMSLPSNGELDPVLE 321
 250 IEVAPDRRLTAEINRHPWTAGCKELELPMMEV-----VQTHVPTAVAPVDVLAN 304
 322 SMAISGCFRERRELRHRELSSEENOEKMIYLLDKREKPSCEDD---LPPRD---V 375
 305 AICISGCFKEKELIOLLESSNTEKVIYELLERRRRPALEDDDELAKSRSELDAY 364
 376 DPPRRKVDSPMLSRHKKRRPERKSMELSTIDAGGSGSPV-PTRRALVMAQS---OHS 430
 365 DPPRRKLDTCRI--NGTAPSYGOL-----SESPILTPRQAFNFRSYSTENHOR 413
 431 RVSASGASGLSSPLSPSP--RSPVFS-----ESPEDA-----GDEARGG-GS 470
 414 RSPTYVTSVRSSTSHPTRCNSPMSAQQOAMASRPSPACTRHSSTYGDNRSGHHS 473
 471 PSEKOTLPS-----RGRGGAGGDEPPPSASTPLPG-PPG 507
 474 SVRFTSHSOKSIGEDVYVVRERPRERDSLKORGGG-----SPRGGCGILPG 525
 508 SP--RSGGCTPLSPPLH-----TPRASP--TGPGTTPPSPGGAGVGAAM 549
 526 SPGGNSGSGTSASPSVHHRANSGLTIAIVNPGSPMANNSSPGM--PGSCNTPGGGLW 583
 550 RSLRNSRNSFLSPRRHRRKMOVPFAEMSLTPESPSELAKRWFONFTSLDEKQIF 609
 584 KTRLTNKNLSLSPRRHRRKMOV-SADEV-FLTPESPSELAKRWFONFTSLDEKQIF 641
 610 IYLKDKPLSSIKADIVAFSLSPSLSHSVLSOTSPFAEYKASG-GPSVQKQVREOVDS 668
 642 ILVKKPLATVAKHLIHLFSLMAELSHSVSPISFVEKFKRNGNGVMEORHKEQVDS 701
 669 SEEGEPSPRRDGGSGGIVSVFTLLSGSPRRKRVETIOALIS 715
 702 AI-----CKGDIADMLFALTFTLLSGNIRRRICEHIOSOVCS 741

RESULT 7
 061298 PRELIMINARY; PRT: 698 AA.

AC 061298; TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE HPOK-1 protein.
 GN HPOK-1.
 OS Halocynthia roretzi (Sea squirt).
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 OC Stolidobranchia; Pyuridae; Halocynthia.
 OC NCBI_TaxID=7729;
 RN
 RP
 RC
 RX MEDLINE=98440280; PubMed=9767157;
 SA Sasaki Y., Ogasawara M., Makabe K.W.;
 RT "Maternally localized RNA encoding a serine/threonine protein kinase
 in the ascidian, Halocynthia roretzi."
 RL Mech. Dev. 76:161-163(1998).
 CC 1. SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AB014885; BAA28663.1;
 DR HSSP: Q63450.1A06
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR002290; Ser_thr_Pkinase.
 DR Pfam: PF00069; Pkinase; 1.
 DR ProDom: PD000001; Euk_Pkinase; 1.
 DR SMART: SM00220; S_TKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE: PS00108; PROTEIN_KINASE_SF; 1.
 KW ATP-binding; Serine/threonine-protein kinase; transferase.
 SO SEQUENCE 698 AA; 78607 MW; B29B1751E8300BF CRC64;
 Query Match 44.2%; Score 1814; DB 5; Length 698;
 Best local similarity 52.8%; Pred. No. 5.6e-112; Indels 118; Caps 21;
 Matches 383; Conservative 109; Mismatches 115;
 25 POHAYVPRLEKTIKGTGKOTGVKGVHCITGOKAIVKREKLSSEVLMKVEREAI 84
 5 POPQOYVPRLEKTIKGTGKOTGVKGVHCITGOKAIVKREKLSSEVLMKVEREAI 64
 85 IKLIEHPVLIADHYENKRYLYLVLEHVSGLPFLYTKKRLPKERKFFQIVSNL 144
 65 KKLIEHPVLIADHYENKRYLYLVLEHVSGLPFLYTKKRLPKERKFFQIVSNL 124
 145 DECHSYSTCHDLPENLLDEKNNIRIADFGMASLQVGDLSLETSCGSPHACPEVING 204
 125 DYCHNNHVCARDLPENLLDEKNNIRIADFGMASLQVGDLSLETSCGSPHACPEVING 184
 205 KYDGRADWMSGVILFALLVGLALPDDNLRKLEKRGVEMHPRFPDCCSLRGM 264
 185 KYDGRADWMSGVILFALLVGLALPDDNLRKLEKRGVEMHPRFPDCCSLRGM 244
 265 KIEVEPERKLSLEOIKHPMT-LOGKHEPD---PCLPAPGRVAMSLPSNGELDPVLE 321
 245 MIDVRPDKRLSLQOYLQHPMWRPGSNVGLVTPDPVAVIDCPV-----LPEES 296
 315 LDPDVLSEMASLGCFFDRERELHRELSSEENOEKMIYLLDKREKPSCEDD---LPPRN 373
 297 VDDPVLASMTSLGCFNCKELKLNLTTEONTEKVIYELLERRRRPALEDDDELAKSRSELDAY 356
 374 DVDPKRRVD--SPMLSRHG-----RPRERK-SMELSTIDAGGSGSPVPTRRALVMAQ 425
 357 HPDAPRRKRVSTSLSSNDDCVNPIPKKMSASLCLTSS--SSPLSRK--KSE 411
 426 HSQRSVYSGASTGLSSPLSPSPSPVSPFGAGDEARGGGSPSTKOTPLPSRPG 485
 412 THORSQSL-----TESSNRL-----VCNISDQTKAESKRNGTTPVRRGTCSS----- 456
 486 GGAGSGPPSPARSTPLPGPSRSGGTPPLSPLTPPASPTGPTTPPSPGGAGV 545
 457 ---NQPVP-----QINTPAS-----FNP----- 471
 546 GAAMSRNLSNLSGSPRRHRRKMOVPFAEMSLTPESPSELAKRWFONFTSLDEKQIF 609
 472 ---WRQRLASLNTGSGSPRRHRRKMOVPFAEMSLTPESPSELAKRWFONFTSLDEKQIF 641
 604 ---KEQIFLYLKDPLSSIKADIVAFSLSPSLSHSVLSOTSPFAEYKASG-GPSV 656
 528 STEHCDELPAIVAKNRLINYSKSELVHAFSLPILTSWVSPTRFCDYSSGSGTSTSV 587
 657 F-QKPVREOVDI-----SSSGPSPRRDGGSGGIVSVFTLLSGSPRRKRVET 708
 588 FHGRSIFQVODIIQHSLSLRQENKKRPSQYGS-----FTIAFSLISGPIRRYRVLEL 642
 709 IOAOL 713
 643 LOMOM 647

RESULT 8
 019469 PRELIMINARY; PRT: 914 AA.

AC 019469; OGBM6;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE F15A2.6 protein (Serine/threonine kinase SMO-1).
 GN F15A2.6.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Rhabditinae; Caenorhabditis.

OX NCBL_TaxID=6239;
 RN [1]
 RA SEQUENCE FROM N.A.
 RA Gregory J.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RA SEQUENCE FROM N.A.
 RA MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode *C. elegans*: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 RP SEQUENCE FROM N.A.
 RA Crump J.G., Zhen M., Jin Y., Bargmann C.;
 RT "The *SAD-1* kinase regulates presynaptic vesicle clustering."
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RC - SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: Z70207; CA94127.2;
 DR EMBL: AF315544; AAG50270.1;
 DR HSSP: O63450; 1A06.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR InterPro: IPR000130; Zn_MTPcdase.
 DR Pfam: PF00069; pkinase.1.
 DR PRINTS: PR00109; pkinase.1.
 DR PRODOM: PD000001; Euk_pkinase.1.
 DR SMART: SM00220; S_krc; 1.
 DR SMART: SM00219; Tykrc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN.1.
 DR PROSITE: PS00108; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN.1.
 DR KAP-binding; kinase; Serine/threonine-protein kinase; Transferase.
 KW SEQUENCE 914 AA; 100840 MW; FDE31D31B249D3E CnC64;
 SQ
 Query Match 42.5%; Score 1744; DB 5; Length 914;
 Best Local Similarity 50.6%; Pred. No. 3.4e-107;
 Matches 392; Conservative 88; Mismatches 198; Indels 96; Gaps 16;

24 PPHAOYVGRYRLEKTLGKQGTGLVLCYHCTGCRVAIKYVREKLSVYMKVERIA 83
 37 PVAOAYCPYRLEKTLGKQGTGLVLCYHCTGCRVAIKYVREKLSVYMKVERIA 83
 84 ILLKLEPHVNLKHYVYENKKTLYLVLEHVSGLGELFDYLVKGRILMSKARKFRQIISA 156
 97 IKLLEPHVNLKHYVYENKKTLYLVLEHVSGLGELFDYLVKGRILMSKARKFRQIISA 156
 144 LDFGHSYICHRDLKPEMLLDKRNIRIADGMAASLOVDSLETSKSGSPHYACPEVIR 203
 157 LDFCHAHNICHRLKPEMLLDKRNIRIADGMAASLOVDSLETSKSGSPHYACPEVIR 203
 204 GEKYDGRADWMSGVILFALLVGLPDDNLRLEKVRGVEHMPHTIPDCCSLR 263
 217 GEKYDGRADWMSGVILFALLVGLPDDNLRLEKVRGVEHMPHTIPDCCSLR 263
 264 GMEVEERKLSLEQLOKHWPYLGKHEPDCLEPAPGRVAMRSLPSSNGELDPVLES 323
 277 AMLEVPGRKYSIADVEFKHPW-VSGTTKADPELELPMSOVYVTHVTPGDSIDPVLRLM 335
 324 ASLGCPRDERELRELRSEENOEKMTYLLDRKERYSCEDQ-DLPPR---NDVDP 379
 336 NCIGCFKDKQKILNELSKNHTKMYVFLDLRRRRRAQDDPEIYLRGAQAQNDP 395
 380 KAVDPSMLSRHGKRRPERKSMVELSITVAGGSGSVPRRLAEMAQHSQORSVSGASTG 439
 396 KRTDSSRTSY-----PMGSIAD-----GSPINRKYTGGRQKSGRHSISLG 437
 440 LSSSPSSRSPV-----FSFSPERAGDEARGGSGSPTEKT-----QTLPS 480
 438 ---SPTESPASSTDLFGSSSSSGSYSARAGEDRDGRGSAKSTNHYHYTPVDPOTLAE 494

QY 481 -----RGRGGGAGE-----QPPPSARSTPLPSPSPNS 511
 DB 495 AARHVRAQERRESRSGSSSRKESKRSKDSKASASSCKNDASSSTVSPHKYSPPSVMS 554
 QY 512 SGGTPLASPLATPRASPPTGTPPTPPSPGGVGGAAAMSRSLNINSLGSPRRHRKM 571
 DB 555 ESVVSSSTWNSTNSNSTSLIAGNSQTSIGSTSG--PMRSKLNINKNSFLGTPFRHRKM 612
 QY 572 QVPTAEEMSLPPE--SSPLAKRSPGNF---ISLDEKQJFLVLDKPLSLIKADYHA 627
 DB 613 SNGTAESEDSQMDITDVLAKSWFGLASSMSYERDTHCVPGKTLNIAELIRA 672
 QY 628 FLSISLSHSVLSQTSFRAEK--ASGSPYFQKPRVQVYISSEGEPSPRRD--SG 684
 DB 673 FLQHELHSHVVGQCFVEYKRGPTVGSVFSKIKANVDIT-----PSPQOVVAGE 726
 QY 685 GGIYSVPTLLSGPSRRKRVETIQADLSTHDQPSVQALADEKNGA--QTRP 736
 DB 727 TPYVVOFVLLAGPVRRFKRLVHLSAIL-----QNSTQQRADRQQAALMVRP 775
 RESULT 9
 ID 095782 PRELIMINARY; PRT; 701 AA.
 AC 095782;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE GH1304P.
 GN CG6114.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Ephygryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OX NCBL_TaxID=7227;
 RN [1]
 RA SEQUENCE FROM N.A.
 RA STRAIN=BERKELEY.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Gonzalez M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Nuncio J., Pacלב D., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celisner S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY060288; AAL25327.1;
 DR FlyBase: FBgn0036544; CG6114.
 DR InterPro: IPR000719; Euk_pkinase.
 DR Pfam: PF00069; pkinase.1.
 DR PRODOM: PD000001; Euk_pkinase.1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN.1.
 DR PROSITE: PS00108; PROTEIN_KINASE_DOM; 1.
 KW KAP-binding; kinase; Serine/threonine-protein kinase; Transferase.
 KW SEQUENCE 701 AA; 76973 MW; 87B308798832425 CnC64;
 SQ
 Query Match 29.9%; Score 1228; DB 5; Length 701;
 Best Local Similarity 45.2%; Pred. No. 3.7e-73;
 Matches 281; Conservative 86; Mismatches 142; Indels 112; Gaps 22;

177 MASLOVGSLLTSCGSPHYACPEVIRGKEDYDGRADWMSGVILFALLVGLPDDNLR 236
 1 MASLOVGSLLTSCGSPHYACPEVIRGKEDYDGRADWMSGVILFALLVGLPDDNLR 236
 237 ROLLEKVRGVFHPHPIPPDCOSLNGMTEVEERKLSLEQLOKHWPY-LGKHPD-- 293
 DB 61 ROLLEKVRGVFHPHPIPPDCOSLNGMTEVEERKLSLEQLOKHWPY-LGKHPD-- 293
 QY 294 -PLEPAPGRVAMRSLPSSNGELDPVLESMAISGCFRDERELRELRSEENOEKMTY 352
 DB 121 LPMKEV-----VQTHYIPATAVDPVLAIGSLGCFKKEKELIDELSSNTEKVIYF 175
 QY 353 LILDRKERYPSCDDQ---LPPRND---VPPRKRVDSPLSRHGKRRPERKSMVELSIT 406
 DB 176 LILDRKRRRPALEDDEDAQKRSSEIDAVDPKRLDLCRI--NCTNAPSTGQI----- 227

```

OY 407 DAGGGSPV-PTRRALM-----AQSQRS-----RSVSGASTGLSSPLS-----447
DB 228 ---SSGSPLEPRQAQFNFRSYSTRHQRSPYTVTSVSRSSSYHSPTRCNSSMAAQO 284
OY 448 -----PRSPV-----PSFSPSP-----AGD-----463
DB 285 ANAISRPSPPAAGTRHSTYGDNRSGHSSVSTPHSHSOKSIEGDVYVVRERIERDS 344
OY 464 --EARGGSPSTKOT-TPSRGPRGGAGROPPPARSTPLPGPGSPRSSGCTPLHSP 520
DB 345 LQERGGGSPRDRDGGCIRPGSGGSSGTSASPSVHHRANSGP-----TIAISMHPD 399
OY 521 LHTPRASPTGTP---GTPP--PPSPGGGAGGAMRSLNRSIRNLSFGSPRHRKMOVPT 575
DB 400 DSNAYVNPNGSPMNNSSPMPSPCNPPTGGQIMTKTKLNKNSFLSPFRHRKMOV-S 458
OY 576 AEMSSLTTPSSPELAKRSFNGFNTSIDKEEQTFLVLDKPLSLKADIVAHALSLPSLS 635
DB 459 ADEY-HITPSSSELPKRSFNGFNLITENKDETFTLTKGKPIATVKALHIAFLSMALSL 517
OY 636 HSYLSQTSFPAEYKAGS-GPSVFOKPYRFOVDISSSEGPSPRRDSSGGGCIYSVFTL 694
DB 518 HSYVSPSPFVEXKRNNGVMEQRHVKFYVDISAI-----CKQGDIALDMLFALFTL 570
OY 695 ISGPRRRKRVETIOQLS 715
DB 571 LSGNIRRRRICEHISQVCS 591

```

RESULT 10

OBSKS5 PRELIMINARY; PRT; 833 AA.

```

ID 08SSKS5
AC 08SSKS5;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative serine/threonine protein kinase.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AK;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
  Lehmann R., Baumgart C., Parra G., April J.F., Gulgo R., Kumpf K.,
  Tunga B., Cox E., Oueli M.A., Platzer M., Rosenthal A., Neigel A.A.;
  "Sequence and Analysis of Chromosome 2 of Dictyostelium."
RT Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AC115685; AAL92711.1; -.
DR KJ.
SQ SEQUENCE 833 AA; 93467 MW; 51B64036C90C0F94 CRC64;

```

Query Match 21.3%; Score 874; DB 5; Length 833;
 Best Local Similarity 50.1%; Pred. No. 1.3e-49;
 Matches 168; Conservative 67; Mismatches 93; Indels 7; Gaps 4;

```

OY 31 VGPYRLKTLGKQGLVGLVHCITGQVAKIVNREKL--SESLMKREIRLAIKLI 88
DB 6 VGFITIGKTLGQGTGKVKIGFHKDTGFKVIGITINKELINRPSMKRIEREIVIMKLI 65
OY 89 EHPHYAKLDVYNNKYLYLYLEHVSGLGELFDLVKKGRLLTPKRAKFFQIYALDFCH 148
DB 66 DHRNNAKMEYVTSKYLFLLEIVBEGELFDLVKGGELGEGEALFFPOIILGLFYCH 125
OY 149 SYEICHKDLKPEMLLDKNNIRIADFGMSLOVGSLSLETSCGSPHYACPEVYKGEYD 208
DB 126 SRNICHKDLKPEMLLDKNNIRIADFGMSLOVGSLSLETSCGSPHYACPEVYKGEYD 185
OY 209 GRADAMSCGVIYLLVGLALPDDNLRQLLEKVRGVFHPHFIPDDCSLHGMIEV 268
DB 186 GOKADVSCGVIYLLVGLALPDDNLRQLLEKVRGVFHPHFIPDDCSLHGMIEV 245
OY 269 EPERLSLEIQKHPWYLGKHEPDLLEPAFGRRVAMRSLPSNGELDDVLESMAISLGC 328

```

```

DB 246 DPKRISIKKIKHPWFVSNFNO---FOKATPYEEINAEPLVDISQIDIFRSIALG- 301
OY 329 FRDRERLHRELSEENQEKMIYLLDRKERYPS 363
DB 302 VGTIDEVKQOLVSNOKS-ATLITRYLLEERKKFDS 335

```

RESULT 11

OBSKS5 PRELIMINARY; PRT; 1371 AA.

```

ID 09Y2K2
AC 09Y2K2;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE KIAA0999 protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA MEDLINE-99246063; PubMed-10231032;
  Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirose M.,
  Miyajima N., Tanaka K., Kotani H., Nomura N., Ohara O.;
  "Prediction of the coding sequences of unidentified human genes. XIII.
  The complete sequences of 100 new cDNA clones from brain which code
  for large proteins in vitro."
RT DNA Res. 6:63-70(1999).
RL -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB023216; BAA76843.2; -.
DR HSBP; Q63450; IAO6.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF00069; pkinase.1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase.1.
DR SMART; SM00220; S_TKC.1.
DR SMART; SM00219; S_TKC.1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Serine/threonine protein kinase.
FT NON_TER
SQ SEQUENCE 1371 AA; 149525 MW; FOFBA385B8226158 CRC64;

```

Query Match 20.3%; Score 832; DB 4; Length 1371;
 Best Local Similarity 28.6%; Pred. No. 1.5e-46;
 Matches 246; Conservative 115; Mismatches 280; Indels 220; Gaps 25;

```

OY 2 SSGAKEGGGSPATVHPH-----HHHPQHQA-----YVGYRLKLT 39
DB 62 AAGAGTGGAGPAGRLPPAPGSPAPAPVSPAAGPAPAPAPAPAPAPAPAPAPAPAPAP 121
OY 40 LGKQOTGLVGLVGVICGQVAKIVNREKLSVLMKVEREIALKLIHPHVLKLDV 99
DB 122 IKGKNEFVVRATHLVYKAKVAIKIIDKTQDLENLKKIFREYQIMKMLCHPHIIRLYOV 181
OY 100 YENKRYLYLVLEHVSGLGELFDLVKKGRLLTPKRAKFFQIYALDFCH 159
DB 182 METFERMAYLVREYVSGGELFDLVKAGHMAKERRKQIVTAVYFCHRNIVHDLKA 241
OY 160 ENLLDKNNIRIADFGMSLOVGSLSLETSCGSPHYACPEVYKGEYD 219
DB 242 ENLLDKNNIRIADFGMSLOVGSLSLETSCGSPHYACPEVYKGEYD 301
OY 220 ILFALLVGLALPDDNLRQLLEKVRGVFHPHFIPDDCSLHGMIEV 279
DB 302 VLYVYVGLALPDDNLRQLLEKVRGVFHPHFIPDDCSLHGMIEV 361

```

QY 280 QKHPWLGKHEPDPCLPAPGRVA-MRSLPSNGELDP---DYLESMASLCPRDRERL 335
 DB 362 CKHKMKMG--DADPNND-----RLNACQQLKEERQVDPLNEDVLLMEDGL--DKEQT 413
 QY 336 HRELSEENQEMKIYLLIDRKERYPCSEDOLP--PR-----NDV 375
 DB 414 LQSLSDAYDHSATISLDCDRKH RHKTLLRGLALPSMPRALAFQAPVINOEAQGTAMNI 473
 QY 376 DPPRRKVDSPMLSRGKRPERKSM---VLST-TDAGGGGSPVPRRALEMAQHS----- 427
 DB 474 SVFOVQLIN-----PENQIVPEDGLINDSDGEEPSPEALVRYLSMRHTVGA 533
 QY 428 -QRSRSVGSASTGLSSPLSPSPVFSPEPGAGDEARGGSPSTKQTLPSRGPRG 486
 DB 524 DPTIEMEDLQKLPGFPVNPQAPFLQVAP-----NVNFMNLPLPMQNLQPT 571
 QY 487 GAGEQP-----PPSARSPLPGPPSPRSSGT--PLHSP--LHTRPA--SP--TGTPT 534
 DB 572 GQLEYKESLQPLPQLLNGMPLGRASDGGANILQHAQQLKRPGRPSPLVWTTPAV 631
 QY 535 TPRPPGGGCGAAMRSRLNSTRNSFLGSPRRKMQVPTAEKMSLTPESSPELAKRS 594
 DB 632 -----PAVTPVDESSDGPDPDEAVOSSTY 656
 QY 595 WGFNFISLDKEQFLVLKDKPLSIKADIYHAFSLISHSVLSQTSFRAEYASGSP 654
 DB 657 KDSNTLHLEPTE-----FSPVRFSDGASIAQFAKHLKMGNN 695
 QY 655 S-----VQKPVRFQVDISSEGPSPRRDGGGGIYVTTLLSGPRRRKR 704
 DB 696 SSIKQLOECERQLOKMYGGQID-----ER 719
 QY 705 VETIOQLSTHDOPSVQALADENGAQOTRPAAP-----RSLQ----- 745
 DB 720 TLETKQOQHMLYQOEQHQLIQOQIIDSICPPSPPLQACENOPALLITLQLOLRTP 779
 QY 746 --PPGRRDPELSSSPRGPP 764
 DB 780 SSPPPHNPNNHLPQSPNSPP 800

RESULT 12
 Q96RCO PRELIMINARY; PRT; 691 AA.
 AC Q96RCO;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
 DE Ser/Thr protein kinase PAK-1beta.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sun T.O., Lu B., Feng J.J., Reinhard C., Jan Y.N., Faull W.J.,
 RT "PAK-1 is a Dishevelled-associated kinase and a positive regulator of
 RT Wnt signaling.";
 RL Nat. Cell Biol. 3:0-0(2001).
 DR EMBL; AF387638; AAK82368.1; -;
 DR InterPro: IPR000719; Euk_kinase.
 DR InterPro: IPR001772; Kinase_Cterm.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR00449; UBA_domain.
 DR Pfam; PF02149; KAI; 1.
 DR Pfam; PF00069; kinase; 1.
 DR Pfam; PF00627; UBA; 1.
 DR ProDom: PD000001; Euk_kinase; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.

KW ATP-binding; Kinase; Transferase.
 SQ SEQUENCE 691 AA; 77631 MW; 2216B4A7BC931BE CRC64;
 Query Match 20.28; Score 828; DB 4; Length 691;
 Best local Similarity 30.94; Pred. No. 1,le-46;
 Matches 232; Conservative 115; Mismatches 261; Indels 142; Gaps 24;
 QY 30 YGPPRLKTLGKQOTGLVGVICIRGQVAKIVRRELSSEVLMKVEREIAILKLE 89
 DB 16 HIGNYRLKLTIGNGNPAKVLARHILTGKEVAKIIDKTOLNSSSLQKLEVRIMKVLN 75
 QY 90 HPAVAKIDVYENKKTLYLVLEHVGSELFDYLVKGRTPKRAKFFQIVASALDCHS 149
 DB 76 HPAVAKIDVYENKKTLYLVLEHVGSELFDYLVKGRTPKRAKFFQIVASALDCHS 135
 QY 150 YSICHDLPKENTLLDEKNRIADFGMSLOYGDSILETSCSPPIYACEVYKGEYDG 209
 DB 136 KEIVHNDIAKENTLLDADNLIKADFGSNEFTFGKLDFTGSPYAAPELFQKKYDG 195
 QY 210 RRADMMSCGVILFALLVGLPFDDNMLKOLLEKVKGFHMPHFIPDDOSILRGMEVE 269
 DB 196 PEVDVMSLVIIYTLVSGSLPFDGQNLKELREVLNKGKRIPEYMETDLENLKLKTLIN 255
 QY 270 PEKRLSLQIOKHPWYLGKHEPD---PCLPAPGRVAMRSLPSNGELDPVLESMASTL 326
 DB 256 PSKRGTLQIMDKRMANNV--HEDDELKPYVEPLPYDK-----DPRTELAVSM 302
 QY 327 GCFRDERELHRELSEENQEMKIYLLIDRKERYPCSEDOLPPLRNDVDPKRVKVDSP- 385
 DB 303 G--YREETQDLSVQORVY--EVMAIYLLLGYSSELBGTITLKRPPADLNTSSAPSPS 359
 QY 386 -MLSRGKRPRKSMKMEVLSTIDAGGSGVPYTRALE---MAOSORSVSGASTGLS 441
 DB 360 HKVQASVSNAPKQRF-----SDQAPALPTSNYSYKKTQSNNAENKRPEDRESGRK 412
 QY 442 SSPLSPSPVFSPEPGAGDEARGGSPSTKQTLPSRGPRGAGAEOPPPSPASSTP 501
 DB 413 AS--STAKVPA--SPLPGL--ERKKTTPPTSTNVSLSTSN-----NSRNSP 453
 QY 502 L--PQPPSPRSSSGGTPLASPLHTRPASPT-----GTPGT----- 535
 DB 454 LLERASLGQASIQNGKSTAPQVAVPSPASHNISSGAGDPRTNPPRGVSSRSTFHAGQ 513
 QY 536 -----PPSPGGGCGAAMRSRLNSTRNSFLGSPRRKMQVPTAEKMSLTPESSPELAKRS 594
 DB 514 LROYDQONLPGVTAPSPSGHSGR--KGASGSIKSF--TSKFEVRNMLNEPESKDRVE 569
 QY 581 SLTP-----ESPELAKRW--FGNFTSLDKEQIFVLKDKPLSITKAD 623
 DB 570 TLRHVVGSGGNDKEKEEFREAKPRSLFTWSMKTTSSMEPNEMRELRYLDANSQSE 629
 QY 624 IVAHLSIPLSHSVLSQTSFRAEYKASGSPVQKRVRYQVDISSEGGPPSPRRDGG 683
 DB 630 LHEKYMML--CMGTGHEP-----VQWEMEYCK-----LPRLSING 665
 QY 684 GGGIYVTFLLISGSPRRKRVVETIOQL 713
 DB 666 -----VREFRISGTSMAFKNISKATANEL 689

RESULT 13
 Q15524 PRELIMINARY; PRT; 745 AA.
 AC Q15524;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-FEB-1997 (Tremblrel. 02, last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
 DE Serine/threonine protein kinase.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_Taxid=9606;
 RN [1]


```

QY 377 -----PERRKVD-----PMLSRHK-----RPE-----RKM-----400
Db 360 HKVORSANPKORRFEDQAAPAIPTNSYSKKTOSNNAENKREEDREGSKAKSIAK 419
QY 401 -----EVLSTIDAGGGSPVPRRLALEMA--OHSGRSRSVSG--435
Db 420 VPASPLPLGLERKATTPPTPSINVSILSTNSRNSPLERKSLGOASLONKDSILMPGSR 479
QY 436 ASTGLSSPLSPR-----SPVSEFSPGACDEARGGSPSTKOTLPSPGP-----483
Db 480 AASTASAAVSAARPROHOKMSASVHPNKAASGLPPRESNCEVPRPTAPPRVVASPSA 539
QY 484 ----RGGAGEO----PPPRASSTPLPGPPSGSPSSGCTPLHSLPTPRASPTGTPTTP 536
Db 540 HNISSGGAPDRTNPRGVSSRSTFHAGOLROYDQONLBY-----CVT- 583
QY 537 PPSPGGVGAAMRSRLNSTRNSFLG-----SPRHRKMOVPTADE--MSLTP-----584
Db 584 PASPGHSGR--RGASGSLFSKFTSKFVARNLSEFRARNLNEPESKDRVETLRPHVG 641
QY 585 -----ESSPELAKRSW--FGNFTSLDKEQIFLVLDKPLSLIKADIYHAFISI 631
Db 642 SGGNDKEKEEFREAKFRSLFTWSMKTSSMEPNEMAREIRKVLNANCSSELHEKTM 701
QY 632 PLSHSHVLSQTSFRAEYKASGSPVQKPVPRQVODISSSGPEPPSRDGGGGGIYSVT 691
Db 702 --CMHGTFGHEDF-----VQWEMEVCK-----LPRLSLNG-----VR 731
QY 692 FTLSGFSRRKRVVETIOAOL 713
Db 732 FKRISTGMAFKINASKIANEL 753
RESULT 15
Q9BYD8 PRELIMINARY; PRT; 688 AA.
Q9BYD8;
AC 01-JUN-2001 (Tremblere, 17, Created)
DT 01-JUN-2001 (Tremblere, 17, last sequence update)
DT 01-JUN-2002 (Tremblere, 21, last annotation update)
DE MAP/microtubule affinity-regulating kinase like 1.
OS MARKL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE=21226021; PubMed=11326310;
RA Kato T., Satoh S., Okabe H., Kitahara O., Ono K., Kihara C.,
RA Tanaka T., Tsunoda T., Yamaoka Y., Nakamura Y., Furukawa Y.,
RA Isolation of a novel human gene, MARKL, homologous to MARK3 and its
RA involvement in hepatocellular carcinogenesis.
RT Neoplasia 3:4-9(2001).
CC -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB049127; BAB39380.1;
DR HSSP; O63450; 1A06.
DR InterPro; IPR000719; Euk_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF00069; kinase; 1.
DR Pfam; PF00627; UBA; 1.
DR PRINTS; PRO109; TYRKINASE.
DR ProDom; PD000001; Euk_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYRKc; 1.
DR SMART; SM00165; UBA; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_SF; 1.

```

```

KW ATP-binding; Kinase, Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 688 AA; 75261 MW; A03B5A7943ACD086 CRC64;
Query Match 20.1%; Score 826; DB 4; Length 688;
Best Local Similarity 32.4%; Pred. No. 1,5e-46;
Matches 235; Conservative 93; Mismatches 223; Indels 174; Gaps 20;
QY 25 PQHAQVGPYRLEKTLGKQGTGLVGLVHCITGOKVAKIKVINEKLSSEYLMKVEREIAI 84
Db 50 PEEQAPVGVNRLRLFTIGKNSAKVLAHLITGREAVALKIIDTQNLPSLQKLFREYRI 109
QY 85 LKIEHPYKLHDYENKRYLYLYEYHNSGGELFDYLVKGLTPREARKFRQIYSAI 144
Db 110 MGLNHPNIVKLFLEYETETLYLVEYASAGEVFDYLVSHGMEKEKARAKFRQIYSAV 169
QY 145 DECHSYSLCHRDLPENLLDEKNIRIDFGMASIQVGSILETSGSPHYACPEYIK 204
Db 170 HCHQKNIVHRLDKAENLLDEANIKIDFGSNEFTLGSKLDTFCGSPYVAPLEFQG 229
QY 205 EKYDGRADMGCGYTLFALLVGLPFDNDNLQELLEKYRGVFNHPHPTPDCQSLRG 264
Db 230 KYDPEVDIMSLGVLYLVLSGLPFDGHNKELREVLKGVYRVFYVSTDCESILIR 289
QY 265 MIEVEPERLSLEQIOKHPWYLGKHEPDPCLEPARGVYAKRSLPSNELPDVYESMA 324
Db 290 FLVLPNAPAKRCLTEQIMKOKWNIYGE-----GEBLKPYTEPEDFGDTKRIEVAV 339
QY 325 SLGCFRDRERLHRELSEENQKMYLLDRKERYPSCEDODLPFRNDVDPKRYVS 384
Db 340 GMG--YTREEKESLTSOKYN--EYATVLLGRKE---EGGD---RAPGLALARYVA 389
QY 385 PMLSRHGKRRPERKSMELVSTIDAGGGSPVPTRALALEMAHOSRSYSGASTGLSSP 444
Db 390 PSDTTNG-----TSSKGTSHSGQSSSTYHRORRHS---DECGPPAP 432
QY 445 LSSPRVSFSPSPGADDEARGGSPSTKOTLPSPRGAGAGEOPPPPARSTPLPG 504
Db 433 LHPKRSPTST-----GEALKEERLPGR-----RASCSTAGSG 465
QY 505 PPSGRRSGGCTPLHSLPTP-----RASPTGTPTTPPPSPGGVGAAMRSRLNSI 556
Db 466 SRGLPSS--PMVSAHPNKAETPERKDSSTPNMLP-----SMATR 508
QY 557 RNSFLGSPR-FHRRKMOVPAEEMSSILPESSPELAKRSWFGNFTSLDKEQIFLYLKD 615
Db 509 RNTYVCTERBGAERPSLLPNKENSSTPRYP----- 541
QY 616 PLSIKADIYHAFILSTLSHVSLSQTSFRAEYKASGSPVQKPVPRQVDSISSEGP 675
Db 542 -----ASPS-SHSLAPPSGERS--RLARGSTI-----RSIFHGQV 574
QY 676 SPRRDGGGGGIYSVFTLISGFSRRKRVVETIOAOLLSTHDPVQALADENGAQTR 735
Db 575 RDRRAGGGGG-----GYONGP-----PASPTLAHE--AAVL 604
QY 736 PAGAP 740
Db 605 PAGRP 609

```

Search completed: April 16, 2003, 12:32:07
Job time : 104 secs

GenCore version 5.1.4-P5-A578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2003, 12:22:22 ; Search time 28 Seconds
(Without alignments)
1152.449 Million cell updates/sec

Title: US-10-003-690-2

Perfect score: 4108
Sequence: 1 MSSGAKREGGSPVHLPHP.....PRGPPKDKKLATNGTLP 778

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	836	20.4	1518	1	KKK1_YEAST
2	816	19.9	786	1	SNIL_HUMAN
3	797	19.4	713	1	PE78_HUMAN
4	795.5	19.4	512	1	K110_ARATH
5	788	19.2	774	1	KEMK_MOUSE
6	784	19.1	776	1	SNIL_MOUSE
7	770.5	18.8	1142	1	SNIL_MOUSE
8	760.5	18.5	576	1	SNIL_MOUSE
9	757.5	18.3	619	1	SNIL_MOUSE
10	753.5	18.3	619	1	SNIL_MOUSE
11	730	17.8	619	1	SNIL_MOUSE
12	725	17.6	550	1	AAK1_HUMAN
13	722	17.6	548	1	AAK1_HUMAN
14	722	17.6	915	1	SNIL_MOUSE
15	720	17.5	611	1	SNIL_MOUSE
16	718	17.5	552	1	AAK2_HUMAN
17	715	17.4	620	1	AAK2_HUMAN
18	710.5	17.3	552	1	AAK2_HUMAN
19	672	16.4	891	1	KIN1_MOUSE
20	666.5	16.2	502	1	KIN1_MOUSE
21	638.5	15.5	622	1	YNA3_MOUSE
22	637.5	15.4	661	1	YNA3_MOUSE
23	617.5	15.0	593	1	CDL1_MOUSE
24	594.5	14.5	1064	1	HUNK_HUMAN
25	570.5	13.9	714	1	HUNK_HUMAN
26	569	13.9	800	1	KIN2_MOUSE
27	556.5	13.5	1147	1	KIN2_MOUSE
28	555.5	13.5	714	1	HUNK_MOUSE
29	550	13.4	664	1	KCCB_HUMAN
30	521	12.7	1050	1	ULK1_HUMAN
31	518.5	12.6	353	1	ASR2_ARATH
32	497	12.1	499	1	KCCD_HUMAN
33	495	12.0	533	1	KCCD_MOUSE

34	482.5	11.7	370	1	KCC1_HUMAN
35	482.5	11.7	1051	1	ULK1_MOUSE
36	479.5	11.7	542	1	KCCB_MOUSE
37	479.5	11.7	542	1	KCCB_MOUSE
38	479.5	11.7	295	1	KMIC_MOUSE
39	478.5	11.6	374	1	KCC1_MOUSE
40	478.5	11.6	478	1	KCCB_HUMAN
41	478.5	11.6	478	1	KCCB_MOUSE
42	477.5	11.6	472	1	KCCB_HUMAN
43	477.5	11.6	363	1	ASR1_MOUSE
44	474.5	11.6	406	1	KPBH_MOUSE
45	471.5	11.5	527	1	KCCG_MOUSE

ALIGNMENTS

RESULT 1
ID KKK1_YEAST STANDARD: PRT; 1518 AA.
AC P34244;
DE 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DF 30-MAY-2000 (Rel. 39, Last annotation update)
DE Probable serine/threonine-protein kinase YKL101W (EC 2.7.1.-).
GN YKL101W OR YKL453.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_Taxid:4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN: S288C;
RX MEDLINE:9407677; PubMed:8256524;
RA Pallier C., Valens M., Puzos V., Fukuhara H., Cheret G., Sor F.,
RA Boloite-Fukuhara M.;
RT "DNA sequence analysis of a 17 kb fragment of yeast chromosome XI
RT physically localizes the MRB1 gene and reveals eight new open reading
RT frames, including a homologue of the KIN1/KIN2 and SNL1 protein
RT kinases."
RL Yeast 9:1149-1155(1993).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC NIM1 SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: X71133; CAAS0456.1;
CC EMBL: Z28101; CAAB1941.1;
CC PIR: S37928; S37928.
CC PIR: S39084; S39084.
CC HSP: 063450; 1A06.
CC SCD: S0001584; YKL101W.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR002290; Ser_thr_pkinase.
CC Pfam: PF00069; pkinase; 1.
CC ProDom: PD000001; Euk_pkinase; 1.
CC SMART: SM00220; S_TKc; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE: PS00110; PROTEIN_KINASE_DW; 1.
CC Hypothetical protein; Transferase; Serine/threonine-protein kinase;
CC ATP-binding.
CC
CC DOMAIN 81 369 PROTEIN KINASE.
CC FT ND BIND 87 95 ATP (BY SIMILARITY).
CC FT BINDING 110 110 ATP (BY SIMILARITY).
CC FT ACT_SITE 239 239 BY SIMILARITY.
CC FT SEQUENCE 1518 AA; 169592 MM; 803F84F7531241DD CRC64;

RESULT 2	SNIL_HUMAN	STANDARD:	PRT: 786 AA.
ID	SNIL_HUMAN		
AC	P57059;		
DT	16-OCT-2001	(Rel. 40, Created)	
DT	16-OCT-2001	(Rel. 40, Last sequence update)	
DT	16-OCT-2001	(Rel. 40, Last annotation update)	

```

DE Probable serine/threonine protein kinase SNFLK (EC 2.7.1.-).
OS Homo sapiens (Human)
GN SNFLK.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN NCBI_TaxID=9606; [1]
RX MEDLINE-20289799; PubMed-10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Seda E.,
RA Ohki M., Takagi T., Sekaki Y., Tauden S., Blechschmidt K., Polley A.,
RA Menzel U., Delagrè J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald C., Rump A., Schillabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudof J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sakai T., Nagamine K., Matsuyama S., Antonarakis S.E.,
RA Moschoma S., Shimizu N., Nordström G., Hornischer K., Brandt P.,
RA Schafte M., Schoen O., Desario A., Reichelt J., Kauer G., Blöcker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riessmann L., Dagand E.,
RA Leinweber S., Borzym K., Gardiner K., Mizetic D., Francis F.,
RA Weinrich R., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21." ;
RL Nature 405:311-319(2000).
CC -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC SNFL SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP001751; BAA95536.1; -.
DR HSPG; P24941; IAQ1.
DR GeneW; HGNC:11142; SNFLK.
DR MIM; 605705; -.
DR InterPro; IPR000719; Euk_pkinase.
DR DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_Trc; 1.
DR SMART; SM00165; UBA; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP_1.
DR DR PROSITE; PS00108; PROTEIN_KINASE_ST_1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM_1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding.
KW DOMAIN 27 281
FT NP_BIND 33 41 ATP (BY SIMILARITY).
FT BINDING 56 56 ATP (BY SIMILARITY).
FT ACT_SITE 149 149 BY SIMILARITY.
SQ SEQUENCE 786 AA; 85252 MW; FD44EE7CAFB7CB1A CRC64;
Query Match 19.9%; Score 816; DB 1; Length 786;
Best Local Similarity 30.9%; Pred. No. 6e-26;
Matches 248; Conservative 103; Mismatches 301; Indels 150; Gaps 26;
OY 31 VGPRLKLTGGGTGLVKGVLGVCHITGCKVAIKTVNREKLSSESVLMVEREIALIKLEH 90
DB 24 VGFDIETELTKGPFNAVKLARHRHVTTQVAIKITDTRLDSSMLERIVREVQAMKLNNH 83
OY 91 PHVLKLDHYENKKYYLYLVIEHVSAGELEFDLYVKGRLTPREAKRFRROYASALDFCHSY 150
DB 84 PHILKLVOMETKMTLVIVEFPANNGEMPDYLTNSGHLSENKAARKRWOLISAEYCSDH 143
OY 151 SICRIOLKPEKLLIDEKNNTRIA---DFGASIQVGDSLSLETSCGSHVACPEVIKGEKY 207
DB 144 HIYRDRLKTENTNLDMMDIKLAGTEFGNGFNFKSGEPSTMGSGSPYAAPFEVEBEKEY 203
OY 208 DGRADMGSCVILFALLVGLALPPDDDNQLLEKRYKRGVFHNPHPTPPCCGSILKRMTE 267
:
```

```

Db 204 EBPQDLSLVLYLVYVCGSLPFDGPNLPTLQRYLGEFRIFPFMSQDCESLIRMLV 263
QY 268 VEPERKLSLEQLOKHPWYLGKHEPDCLEPARG-RRVAMRSLPSN-GEIDPDVLESMA 325
Db 264 VDPARITIAQIQHMM-----RAEPCL-PGAPAFSAHSTYSLNGDYDEQALGIMOT 317
QY 326 LGCFRDRERHRLRESEENQEMITYLLDKREKERVSCEDDLPRLNDVDPKRVNDSP 385
Db 318 LGV--DROFTVESLONSSYHNFAYLYLLERKEYNNAO-----355
QY 386 MLRHHG-KRRPERKSMELVITDAGGSGSVPTRRALMAOHQSRSKSVGASTG-LSS 442
Db 356 -CARPAPKPPRRRSLSLELVEQBELSDPFRPALLCPQPTLVQSVLAEDMCELOS 414
QY 443 S-----PLSPRRSVFESFSP-----EPGADDERGGSGSPSKQYLRSPRPGGA 488
Db 415 SLQMPLEFPYDASGSGYFRPRVPSLSLDTALISEARQGLLEEDDTORSL-PSSTGR 473
QY 489 GEQPPRSASSTPLPGP-----PGSPRSSGGTPLH-----518
Db 474 RHTLAEVSTRLSLPLADCKTVSPSTTASPAAGTSSDCLTFSAKSPAGLSGTATQGL 533
QY 519 ---SPLHTPRASP-TGTPGTTPPSPGGCGAAMRSRLNSIFLGSPPFHRKQV 574
Db 534 GACSPVRL--ASPLGSGSATPVLAQGLGAV-----LLPVSF--QEGRRASDTS 581
QY 575 TAEKMSITPSSPELAKRSMFQNFISDKREOFVLVKDKPLSLK--ADIVAFSLI 631
Db 582 LTQGLKAF--ROOLKRTTPTKGLTKTKIGLAKQVCOAPASASGGLSPLFAPAS 637
QY 632 P-----SLHSVLSQTS-FRAEYKASGSPSVFQKPVRFQVYDSSSEGPSP 677
Db 638 PGLGGAAGSREGSLLEVELEQRLQLOHNAAP-----GCSQAPAPAP 684
QY 678 RR-----DSGGGGSIVTPT-----LISGSRKRVVETQADLSTHDPVSQALAD 727
Db 685 APFVAPDGGGAAPLPTLLTSLGRLPRLLOTGASPVASADLDTL-----LH 736
QY 728 EKNGAOTRPAAGAPRSLQPPG 749
Db 737 IGTGPTALPAPVPRPLARLAPG 758

RESULT 3
KP78 HUMAN STANDARD: PRT: 713 AA.
AC P27448:
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE putative serine/threonine-protein kinase P78 (EC 2.7.1.1.-).
GN P78.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Maheshwari K.K., Som S., Parsa I.;
RA Submitted (JAN-1992) to the EMBL/GenBank/DBJ databases.
CC -1- MISCELLANEOUS: MARKER PROTEIN LOST IN CHEMICALLY INDUCED
CC TRANSPLANTABLE CARCINOMA AND PRIMARY CARCINOMA OF HUMAN PANCREAS.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC NIM1 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC -----

```

```

DR EMBL: M80359; AAA59991.1;
DR PIR: S27966; S27966.
DR HSP: 063450; 1A06.
DR Genew: HGNC: 6897; MARK3.
DR MIM: 602678;
DR Interpro: IPR000719; Euk_pkinase.
DR Interpro: IPR001772; Kinase_Cterm.
DR Interpro: IPR002290; Ser_thr_pkinase.
DR Interpro: IPR000449; UBA_domain.
DR Pfam: PR00069; pkinase.1.
DR Pfam: PR00067; UBA.1.
DR Pfam: PF02149; KAL.1.
DR ProDom: PD000001; Euk_pkinase.1.
DR SMART: SM00220; S_tkc.1.
DR SMART: SM00165; UBA.1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST.1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding.
KW DOMAIN
FT NP_BIND 56 307 PROTEIN KINASE.
FT BINDING 62 70 ATP (BY SIMILARITY).
FT ACT_SITE 85 85 ATP (BY SIMILARITY).
FT ACT_SITE 178 178 BY SIMILARITY.
SQ SEQUENCE 713 AA; 79903 MW; 1CA78EB22620A228 CRC64;

Query Match 19.4%; Score 797; DB 1; Length 713;
Best Local Similarity 32.0%; Freq. No. 3.1e-25; Indels 138; Gaps 28;
Matches 237; Conservative 112; Mismatches 254;

QY 30 YGPRLEKTEKLGQGTGLKLVCHCTGQVAKIKYVNRKLEESVLMKVEREITAILKLE 89
Db 52 HIGNRRLKTIKGNFAKNAKLARHILITGREAIIKIDKQLNPTSLQKLFREVRIMKILN 111
QY 90 HPHVLKLDVYENKKYLYLVLEHVSQGELEFDLYLKKGLTPKPAKFFQYISALDFCS 149
Db 112 HPIYVLFVIEVETQKLYLIMEYASGKVFQDYLVANGKKEKARSKFYQISAVQYCHQ 171
QY 150 YSLCHDLKPEMLILDEKNNIRADFMAISLQVDSILETSQSPHYACPEVYIKGEYDG 209
Db 172 KRIYHDLKAEMLLDADNINIKIADFGFSNEFTVGSKLDTFQCSPPYAAPELFGKKYDG 231
QY 210 RRADWMSGVILFALLVGLLPFDDNLRLQLEKVKRGVEMHPIFPDCOSLLRGHIEVE 269
Db 232 PEVDVMSLGIYLYTVLVSGLPFDGQNLKELREKRVLYKRYIPYMSYDLENLRLRYLN 291
QY 270 PEKRLSLEQLOKHPWYLGKHEPD---PCLERAPGRVARSLSLPSNGEL---DPVLES 323
Db 292 PIRKGTLEQIMKORWLNAG-HEEDLKPVEP-----ELDISQKRIDIM 335
QY 324 ASLQCFRDRERHRLRESEENQEMITY-----YLLDRKERYPSCEDDLPRLND 374
Db 336 VGMG-----YSQELIGESLSKMKYDEITATYLLGRK-----SSKVRPSSD 376
QY 375 VD-----PPKRYDSPMLSHRGKRRPERKSMELVITDAGGSGSP---VPTRALMAQ 425
Db 377 LNNSTGSPHHKQVRSVSSQOKRR-----YSDHAGPGIPSVAYPKRSQSTAD 426
QY 426 HSOR-----SRVSGASTGLSSPSLSPRSVFSFSEPGAGDEARGGSPSTQTLPS 480
Db 427 SDLKEDGISRKSGSAVGGK---IAPASMLGNANPNKADLPE-----RKSSVPS 478
QY 481 RGRPGGAGDEQPPRSASSTPLPGPSPSSGGTPLHSLPHTRASPTGTPGTPPSP 540
Db 479 SNTASGCM-----TRRNTYV---SSERTTDR--HSLVQNGKENST-IPDQRTPVAS 524
QY 541 GGGVGGAMRSRL-----NSIRNSFLSPRRHRRMQVPTAEKSSSLPSESPELAKRS- 594
Db 525 THSISSAATPRKRPKPTASRSTFHGQPR-ERKATATNGPAPSLSHKATPLSQSTR 583
QY 595 ---WFGNFTS-LDKREQIFVLKD---KPLS-----SLK-----ADIVAFSLIP 632
Db 584 GSTLLFSKLTSLKLRNRYSAKQDENKAKPRSLRFTYMSKTTSSMDPGDMREIRYVL 643

```


Db 309 NHLIESLRNRQNDGTVTYVILLNDRFRASSGYLGAEFQETMEGTPRMHAPESVSPVSH 368
 QY 389 R 389
 Db 369 R 369

RESULT 5
 KEMK_MOUSE STANDARD: PRT: 774 AA.
 ID KEMK_MOUSE 005512;
 AC 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Putative serine/threonine-protein kinase EMK (EC 2.7.1.-).
 GN EMK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=93364122; PubMed=8358177;
 RA Inglis J.D., Lee M., Hill R.E.;
 RT "Emk, a protein kinase with homologs in yeast maps to mouse
 CC chromosome 19.";
 CC Mamm. Genome 4:401-403(1993).
 RL -1. SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC NIM1 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X70764; CAA50040.1; -
 CC HSSP: 063450; 1A06.
 DR MGD: MGI:99638; Emk.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR001772; Kinase_Cterm.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase.1.
 DR Pfam: PF02149; KAI.1.
 DR Pfam: PD000001; Euk_pkinase.1.
 DR SMART: SM00220; S_TKc.1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST.1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 53 304 PROTEIN KINASE.
 FT NP_BIND 59 607 ATP (BY SIMILARITY).
 FT BINDING 82 82 ATP (BY SIMILARITY).
 FT ACT_SITE 175 175 BY SIMILARITY.
 SQ SEQUENCE 774 AA; 85874 MW; 02BF8D7BF443483A CRC64;

Query March 19.2%; Score 788; DB 1; Length 774;
 Best Local Similarity 30.1%; Pred. No. 7.5e-25;
 Matches 238; Conservative 119; Mismatches 261; Indels 172; Gaps 27;

QY 30 YVGYRLEKTLGKGTGLVKGHCITGOKVAIKIVNREKLSVLMKVEREIALIKLIE 89
 Db 49 HIGNRRLKLTIGKGNFAVKIARHLITGKEVAVKIIDTKQLNSSSIQKLFREVRIMKVIN 108
 QY 90 HHVIAKLIDVYNNKKYIVLVEHVGSGLEDFLVKGRGLTPKRAKFFQIYSAIDFCHS 149
 Db 109 HHVIAKLIDVYNNKKYIVLVEHVGSGLEDFLVKGRGLTPKRAKFFQIYSAIDFCHS 149
 QY 150 YSICRDLKPEMLIDLEKNNIRIADFGNALSIVGDSLETSCGSPHYACPEVIGKEKYUG 209
 Db 150 YSICRDLKPEMLIDLEKNNIRIADFGNALSIVGDSLETSCGSPHYACPEVIGKEKYUG 209

Db 169 KPIVHDLAENILLADANIKIADFGSNEFTFGNKLDTFFCSPPYAAPELFGKKIDG 228
 QY 210 RRADWMSGCVILFALLVGLPFDDNLRQLLEKVKGVFHFHPTIPDCOSLIRGMIEVE 269
 Db 229 PEVDVMSLGYITLVSGSLPFQGNLKELEHRYLRGKRIPIFYMSIDCENLKKFLILN 288
 QY 270 PEKRLSLEQIQKHPVTLGGHNEPDCLEPAPGRVAMRSLPSNGEIDPDVLESASIGCF 329
 Db 289 PSKRGTLEQIMKDRMNVG-HEDD-----ELKPYV-EPLITIGP- 325
 QY 330 RDR-----ERLHRELSEENOEKMIYLLIDRKERYPSCEDDLPDRNDVD- 376
 Db 326 RDVVDVNGHLETHEIDSLVGRYN-EVMATYLLGLKSSPEEDITLKPFRPADLTNS 384
 QY 377 -----PPKRYVS-----PMLSRCK-----RRERKSMELVSI 405
 Db 385 SAPSPSHKVOVSANRKHRRSDPAVPAIPTNSYSEKTKQSNNAENKRDEEETGRASS 444
 QY 406 TDAGGGGSPYP--TRALMAHSGRSRVSAGSTGLS--SSPLSPSPVFSFSPBP- 460
 Db 445 T-AKVPASPLPLGLDKKTTTPA--PTNSVLTSTNRSRNSPLLD-RASLGQANSIQNGKD 499
 QY 461 ----AGDEANGGSPSTKTOTLPSPRGGGAGBOP-----PPSANSTPLPGPGSPRS 511
 Db 500 SLIMPGRASPTASASAAVSAARPHQHKMSGASVHPMKASGLPPTESNCEVPRPSTAPOR 559
 QY 512 SGTPLHSPHTPPASPTPTGTT-----PPPSG 541
 Db 560 --VPVAPSPAHNISSSGADRTNPPRGVSSRSTFHAGQLRQYRDOQNLPGYVTPASPS 616
 QY 542 GGVGAAMRSRLNIRNSFLSPFHRKMQVPTAE-MSSLTP----- 584
 Db 617 GHSGR--RGASGSLFSNF--TSKFEVRNLNPESEKDKVELTRPHVSGGTDKDEEER 672
 QY 585 ESSEELAKRSY-FGNFISLDEQITFLVKRPLSSIRADYVHAFSLPSLSHVSLSOTS 643
 Db 673 EAKPRSLFTWSMKTSSMEPEMREIRKYLDANSCSELHERYMLL--CVHGTPGHEH 730
 QY 644 FRAEYKASGGSYVQKPVRFQDSSSGPSPRDSGGGIGYVTFLLSGPSRRFK 703
 Db 731 F-----VQMEKVC-----LPLRLSNG-----VAFKRISGSMAFK 762
 QY 704 RVETIQAOQL 713
 Db 763 NIASKIANEL 772

RESULT 6
 SNIL_RAT STANDARD: PRT: 776 AA.
 ID SNIL_RAT 09R105; 09R081;
 AC 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative serine/threonine protein kinase SNF1LK (EC 2.7.1.-) (Salt-
 DE inducible protein kinase) (Protein kinase KID2).
 GN SNF1LK OR SIX OR KID2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Adrenal gland;
 RX MEDLINE=99330184; PubMed=10403390;
 RA Wang Z., Takemori H., Halder S.K., Nonaka Y., Okamoto M.;
 RT "Cloning of a novel kinase (SIX) of the SNF1/AMK family from high
 RT salt diet-created rat adrenal."
 RL FEBS Lett. 453:135-139(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Feldman J.D., Vician L., Crispino M., Hoe W., Baudry M.,
 RA Herschman H.R.;

RT "The kid2 gene encodes a protein kinase induced by depolarization in brain."
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC SNF1 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB020480; BAAB2673.1; -
 DR EMBL; AF106937; AAF14191.1; -
 DR HSSP; P24941; IAQ1.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR000449; UBA_domain.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00165; UBA; 1.
 DR SMART; SM00165; UBA; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 27 278 PROTEIN KINASE.
 FT NP_BIND 33 41 ATP (BY SIMILARITY).
 FT BINDING 56 56 ATP (BY SIMILARITY).
 FT ACT_SITE 149 149 BY SIMILARITY.
 FT CONFLICT 473 473 R -> K (IN REF. 2).
 SO SEQUENCE 776 AA; 84908 MW; 7BF745AE28F1E6E CRC64;
 Query Match 19.1%; Score 784; DB 1; Length 776;
 Best Local Similarity 30.5%; Pred. No. 1.1e-24;
 Matches 240; Conservative 104; Mismatches 257; Indels 186; Gaps 28;
 QY 31 VGFPRLEKTEGKGTGLVKGVCITGOKVAIKIYNRELSSVLMKRETAIIKLIEH 90
 DB 24 VGFPRLEKTEGKGTGLVKGVCITGOKVAIKIYNRELSSVLMKRETAIIKLIEH 90
 QY 91 PNYKLIDVYENKRYLYLVLEHVSGLDEFLVKKGRTPRKARFQIVASALDPCSHY 150
 DB 84 PNYKLIDVYENKRYLYLVLEHVSGLDEFLVKKGRTPRKARFQIVASALDPCSHY 150
 QY 151 SICHRODKPENLLDEKNNIRIADFGNASLQVGSLEFTSCGSPHYACDEVTKGKRYDGR 210
 DB 144 SICHRODKPENLLDEKNNIRIADFGNASLQVGSLEFTSCGSPHYACDEVTKGKRYDGR 210
 QY 211 RADWMSGVILFALLVALGPEDDNLRQLLEKVKRGVFMHFTPPDCOSILKGMIEYEP 270
 DB 204 RADWMSGVILFALLVALGPEDDNLRQLLEKVKRGVFMHFTPPDCOSILKGMIEYEP 270
 QY 271 EKRLSEQLQKHWYLGCKHEPDLCPAPGRVAMSLPSN-GELDDVLESMASLGF 329
 DB 264 EKRLSEQLQKHWYLGCKHEPDLCPAPGRVAMSLPSN-GELDDVLESMASLGF 329
 QY 330 RURERLHRELSEENOEKMIYLLDRKREYPCEDODLPPRDVDPKRVDSPLMSR 389
 DB 318 RURERLHRELSEENOEKMIYLLDRKREYPCEDODLPPRDVDPKRVDSPLMSR 389
 QY 390 HCKRRPERKSMELVSLTDAGGGSPV-----PTRRALEMAQHQSQRSGVGA--STGLS 441
 DB 368 HCKRRPERKSMELVSLTDAGGGSPV-----PTRRALEMAQHQSQRSGVGA--STGLS 441
 QY 442 SS-----PLSSRPVY-----EPGAGDEARGGSGPTSKQTLLPSRPGGGA 488
 DB 414 SS-----PLSSRPVY-----EPGAGDEARGGSGPTSKQTLLPSRPGGGA 488
 QY 489 GROPPPPSARSTPLPSPSPRSGGTPLH-SPLHMP-----RASPT-GTGTTPPP 538
 DB 462 GROPPPPSARSTPLPSPSPRSGGTPLH-SPLHMP-----RASPT-GTGTTPPP 538

DB 463 -----QEPLEGSTGRHRTLAEVSTHESPLNPPCIIVSSAAVSPSEGTSSDCLP 512
 QY 539 -----SPGGVGGAAMRSRLNS-----IRNSFGSPREHRRKQVPAEMSSITPSSP 588
 DB 513 FSASEGPAGLGGGLATPGLGTSSSPVRLASPFGLS-----QATVYLQAGLGLATLP 566
 QY 589 -----ELAKRSWRCNFTSLDKEQIFVLKDKPLSIKADI 624
 DB 567 PVSFQEGRRASDTSLTGKIAKARQOILKRNARKTKGFLGLNK----- 606
 QY 625 VHAFLSIPSLSHSVLSQTFRAEYKASGSPVQKRYFRQVDISSEGGPPSRDGGSG 684
 DB 607 -----INGLANQV-QQSSIRG-----SRGMSITFHHPA-----PSSGLQG-- 640
 QY 665 GGIYSVTFTLLSGPSRRKRYVETI--QAOLLSTHDPQVQALADEKNGAQTFRPAGAPR 742
 DB 641 -----CTASSREGRSLLEEVYHQORLLQLQHSNVS--SDVQAQAPLSPV--PY 685
 QY 743 SLOPPG 749
 DB 686 VLRPDG 692
 RESULT 7
 ID SNIL_MOUSE STANDARD; PRT; 779 AA.
 AC 060670;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable serine/threonine protein kinase SNF1LK (EC 2.7.1.1-) (HRT-20)
 GN (Myocardial SNF1-like kinase).
 GN SNF1LK OR MSK.
 OS Mus musculus. (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND REVISIONS TO 8; 16-18; 44-45; 316 AND 435.
 RC TISSUE-Embryo;
 RA Ruiz J.C.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-435 FROM N.A., AND TISSUE SPECIFICITY.
 RC TISSUE-Embryo;
 RX MEDLINE=95200798; PubMed=7893599;
 RA Ruiz J.C.; Conlon F.L.; Robertson E.J.;
 RT Identification of novel protein kinases expressed in the myocardium
 RT of the developing mouse heart.*;
 CC Mech. Dev. 48:153-164(1994).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN LUNG, SKIN, OVARY, HEART AND
 CC STOMACH. NO EXPRESSION IN BRAIN, LIVER OR SKELETAL MUSCLE.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC SNF1 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U11944; AAB67926.2; -
 DR HSSP; P24941; IAQ1.
 DR MGD; MGI:104754; Snf1lk.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR InterPro: IPR000449; UBA_domain.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.

DR SMART: SM00220; S.TKC: 1.
 DR SMART: SM00219; TYKIC: 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_DOM; 1.
 DR TRANSFERASE; Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 27 278
 FT NP_BIND 33 41
 FT BINDING 56 56
 FT ACT_SITE 149 149
 FT SEQUENCE 779 AA; 85027 MW; 7808131BC46D9C4E CRC64;

Query Match 18.8%; Score 770.5; DB 1; Length 779;

Best Local Similarity 29.9%; Pred. No. 3.7e-24; Indels 189; Gaps 28;

Matches 235; Conservative 112; Mismatches 251;

31 VGPYRLKLTGKGTGVLKGLVHCITGKVAIKYINREKLSVLMKVEREIAIKLIEH 90
 24 VGFYDVERTIGKGNFVAVKLAHRYKTKQVAIKIIDKTLSDSSNLEKIYREVOLMKLHNH 83
 91 PHVYKLDHYENKKTLYLVLEHVSGLDFDYLVKKGRLLPKARKRFQIVSALDFCHSY 150
 84 PNIKLYOMETKMDLYTEPAKNGEMFDYLSNGHSENEKQFMOILSVEYCHNH 143
 151 SICHRLDKPENTLLDEKNNIRIADFGMASLOYDLSLETSCGSPHYACEVYKGEYDGR 210
 144 HYNHDKLTENILDSNMDIKLADFGFNGFYKPGPELSTCVSPYAAAEVFEKGEYGP 203
 211 RADMSGCVILFALLVGLVFPDDNLRQLLEKVKRQVFMHPIPPDCOSLNGMIEVER 270
 204 QLDVMSIGVLLVYVCGSLPFDGPNLPRTLKORVLEGRFRFPFMSDDCELTIRMLVDP 263
 271 EKRLSLEIOKHFWYVGGHEPDCLEPARGVARSLEPN-GELDPDYLESMASICF 329
 264 AKRTIAQIKRHHM-----QADPTLQDDPDMQGTNSNGVYNGVYLGMAIGI- 317
 330 RDRERLARELSEENQEKMTYLLDRKERYPSCEDQDIPRNDVDPKRRVDSPLMR 389
 318 -DRQRTIESLONSYNHFAIYLLERLEKHSAQ-----PSSRPTAP- 361
 390 HGRKREKSMELVLT-----TDAGGGGSPVPRRALEMAHSGRSVSGA--STG 439
 362 --TRQOLNSSLSSLEVPQELLPCD-----PERSLLCPQALASVLAQAEIDCD 411
 440 LSSS-----PLSSPRSPV---SFSP---EPGAGDEANGSGSPSTQTLPSRPRGG 486
 412 LHSLLQPLFPLDTNCSGVRRHSISPSLLDPAISEANQGPSLEEQEV----- 462
 487 GAGEQPPPSARSTPLPGPGSPSSGGTPLH-SPLHP-----RASPTGP--GTP- 536
 463 -----OEPILGSTGRHTLAVESTHSPPLNPCTIYSSSATASPESEGTSSDSC 510
 537 -----PSPGGGVG--GAAMRSRLNSIRNSFLGSPFRHRRMQVPTAEEMSLPES 586
 511 LPPSASEGPAGLGGALTPGLLGTSSVRLASPLGS-----QSAPVLOTQAGLGNV 564
 587 SP-----ELAKSMFGNFIISLDEKEQIFLVLDKPLSITKA 622
 565 LPPVSPFEGRRASDTSLTGLKAFRQOLRKNAFKGLGLNK----- 606
 623 DIYHAPLSTLSHSVLSQTSFRAEYKASGSPVFOKVRQVYDLSSEGEPPSPRRDS 682
 607 -----IKGLAQV--QSSVRT--PRGMSFTHPA-----PS----- 635
 683 GGGGIVSVFTLLSGSPRRFRVETI--QAQLSTHDPQSVQALADKNGAOTRRPAG 740
 636 --SGLGCTTS-----KREGSLLEEVLAHQRLQLDQHSSTAAS--SCGQCGPQLSP 685
 741 -PRSLQ 746
 686 VPVYLA 692

RESULT 8

GIN4_YEAST

ID GIN4_YEAST

STANDARD; PRT; 1142 AA.

AC 012263;

DT 01-NOV-1997 (Rel. 35. Created)

DT 01-NOV-1997 (Rel. 35. Last sequence update)

DT 30-MAY-2000 (Rel. 39. Last annotation update)

DE Serine/threonine-protein kinase GIN4 (EC 2.7.1.1.-).

GN GIN4 OR YDR507C OR D9719.13.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-99030835; PubMed-9813093;

RA Longline M.S., Fares H., Pringle J.R.;

RT "Role of the yeast Gln4p protein kinase in septin assembly and the

RT relationship between septin assembly and septin function.";

RL J. Cell Biol. 143:719-736(1998).

RL [2]

RP SEQUENCE FROM N.A.

RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E., Berno A.,

RA Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,

RA Hunnicke-Smith S., Hyman R., Komp C., Lashari D., Lew H., Lin D.,

RA Mosedale D., Nakahara K., Nemath A., Oefner P., Oh C., Petel F.X.,

RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,

RA Whant A., Yelton M., Botstein D., Davis R.W.;

RL Submitted (Aug-1995) to the EMBL/Genbank/DBD databases.

CC -1- FUNCTION: MAY PLAY A ROLE IN SEPTIN ASSEMBLY.

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC NIM1 SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: U33140; AAA75513.1; -

CC DR EMBL: U33057; AAB64949.1; -

CC DR HSSP: O63450; 1A06.

CC DR SGD: S0002915; GIN4.

CC DR Interpro: IPR000719; Euk_Pkinase.

CC DR Interpro: IPR002290; Ser_thr_Pkinase.

CC DR Pfam: PF00069; Pkinase; 1.

CC DR Prodom: PD000001; Euk_Pkinase; 1.

CC DR SMART: SM00220; S.TKC: 1.

CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

CC DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

CC DR PROSITE; PS00109; PROTEIN_KINASE_DOM; 1.

CC KW Transferase; Serine/threonine-protein kinase; ATP-binding.

CC FT DOMAIN 27 278

CC FT NP_BIND 33 41

CC FT BINDING 56 56

CC FT ACT_SITE 149 149

CC FT SEQUENCE 779 AA; 85027 MW; EC16F4BB49DD811 CRC64;

SQ

Query Match 18.5%; Score 760.5; DB 1; Length 1142;
 Best Local Similarity 38.3%; Pred. No. 1.3e-23;
 Matches 184; Conservative 77; Mismatches 139; Indels 81; Gaps 13;

31 VGPYRLKLTGKGTGVLKGLVHCITGKVAIKYINREKLS-----ESVL 75
 16 IGPMLIGETLIGSGTKVQLARNGSTGDAVAKVSKAVFNCGVSGTSTPDLAP 75
 76 MKVEREIAIKLIEHPHYKLDHYENKKTLYLVLEHVSGLDFDYLVKKGRLLPKARK 135
 76 VGIEREITIMKLNPNVYKRLDWETNDLYLVLEAKGELFNLVGRPLPEHAI 135
 136 FFRQIVSALDFCHSYSICHRLDKPENTLLDEKNNIRIADFGMASLOYDLSLETSCGSPH 195

```

Db 136 EFROITIGSYCHALGIVHDLKPEMLLDHKNYIKIADFGMALEFEGKLETSQSPH 195
QY 196 VACEYIKGEKYGRADMGSCVILFALLVGLPDP--DDNLRQLEKRYKGFHNP-- 251
Db 196 VAPAEIVSIPVPGFSDVMSCVIIFALLTGRLPFDEEDGNITLLKKGFEFMSD 255
QY 252 HETPPDCGLKMGTEVEPERKLSLEQIOKHPMYLGGKHEPDPCLPEAPGR--VAMRSL 309
Db 256 DEISREADOLIKRILVDEPERIKTRDILKH-----PLLOKYPISROSKSLRGL 304
QY 310 P-----SNGELDDVLESMASLSCGFRDRELRHRLSEENQOEKMYILLDRK 358
Db 305 PREDYITPLSLSNSSIDTITQNLVILMHGRDPEGIKERLEGANAETLVALLYRFK 364
QY 359 ERYPSCE-----DODLPPRNDVD-----PPKRVDSPLMRHGRKRRPERKSMELVITD 407
Db 365 -----CDYQKELIKQOQVKKQSSISVSVPKRV-----STTPQRRRNESL--LSVTS 412
QY 408 AGGGSPPTRALALEMAHQSORSVSGASTGLSSPLSSPSP-----VFSFSEPG 460
Db 413 S-----RKKPISFNKFTASSASSNLTPQSSKRLSKNFSSKKKSLTIVQSSPTPA 464
QY 461 A 461
Db 465 S 465

```

RESULT 9

```

SNFL_SCHPO STANDARD; PRT; 576 AA.
AC 074536;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUN-1999 (Rel. 38, Last sequence update)
DE SNFL-1-like protein kinase (EC 2.7.1.1.).
GN SPEC74.03C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
NCBI_TaxID=4896;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; Pubmed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras K., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jørgels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Moorey P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Glynnoprez B.,
RA Welfens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RA *The genome sequence of Schizosaccharomyces pombe*.
RA Nature 415:871-880(2002).
CC - SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

```

SNFL SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

DR EMBL; AL031543; CAA20833.1; -.
DR HSSP; Q63450; 1A06.
DR InterPro; IPR000719; Euk_Pkinase.
DR InterPro; IPR002290; Ser_thr_Pkinase.
DR InterPro; IPR000449; Uba_domain.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Euk_Pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_ST; 1.
DR TRANSFERASE; Serine/threonine-protein kinase; ATP-binding.
KW DOMAIN 34 285
FT NP_BIND 40 48 ATP (BY SIMILARITY).
FT BINDING 63 63 ATP (BY SIMILARITY).
FT ACT_SITE 156 156 BY SIMILARITY.
SQ SEQUENCE 576 AA; 65996 MW; E5857E8F171E7B50 CRC64;

```

Query Match 18.4%; Score 757.5; DB 1; Length 576;
 Best Local Similarity 36.5%; Pred. No. 9,1e-24;
 Matches 167; Conservative 95; Mismatches 133; Indels 63; Gaps 12;

```

QY 24 PPO--HNOVGYPLEETKIGCGQIVLVGVCITGQVAIVKIVREKLSSEVL-MKYER 80
Db 22 PPEAIRSRHNGEYITREITLGGSGFKVLAHYKQVAKLFISROLKSDMHMRYER 81
QY 81 EIALIKLIEHPVLAHDVYENKKYLVLEHVSNGGEFLVYKGRITLPEAKRFFQOI 140
Db 82 EISYALKLRPHPHIKLVDTITPPDIYVWIEY-AGGELFDYIVKKNTEDEGRFFQOI 140
QY 141 VSALDPCHSISICRDLKPEMLLDKKNIRIADFGMASLOVGDLSLETSCGSPHYAPE 200
Db 141 ICALIEYCHRRKIYHRDLKPEMLLDMLNVAIDFGLSNIMTDGNEFLTSCGSPHYAPE 200
QY 201 VIKGEYGRADMGSCVILFALLVGLPDDDNLRQLEKRYKGFHNPFPDQCS 260
Db 201 VINGKLYAGEVDVMSCVIIFALLVGLRPLPDDDEFIPVLEKRVSCVYVMPDPLSPGAOS 260
QY 261 LKMGTEVEPERKLSLEQIOKHPMYLGGKHEPDPCLPEAPGRVAMRSLPSNGELDPD 318
Db 261 LIRRMIVADPMQRITTOEIRDPWF--NVNLPDLRMEVQGSYADSRVSKIGE----- 314
QY 319 VLESMASLSCGFRDRELRHRLSEENQOEKMYILLDRKERYSCEDOLPPRNDVDP 378
Db 315 -----AMGESD--YVEALRSDENNEVEKAYNL-----HENOVLOEKSHLS-K 356
QY 379 RRRVDS-----PMLSRHGRRRPERKSMELVITDAGCGGS----- 413
Db 357 SKRVDSFLSVSPAFSEYSEIQKSKQKQELIDPLTEBPRTVSDPPYATQOTIDSNICVL 416
QY 414 -VPTPRALALEMAHQSORSVSGASTGLSSPLSSPSPS 450
Db 417 VPTAEKNKLEM-----RTIADAAAVDTSQSTRKKS 447

```

RESULT 10

```

SNFL_YEAST STANDARD; PRT; 633 AA.
AC P06782;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Carbon catabolite derepressing protein kinase (EC 2.7.1.1.-).

```

GN SNF1 OR CARI OR CCR1 OR PAS14 OR GIC2 OR YDR477M OR DB035.20.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86289463; PubMed=3526554;
 RA Celenza J.V., Carlson M.;
 RT "A yeast gene that is essential for release from glucose repression
 encodes a protein kinase.";
 RL Science 233:1175-1180(1986).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94131988; PubMed=7905477;
 RA Mitchellhill K.I., Stapleton D., Gao G., House C., Mitchell B.,
 RA Katsis F., Witters L.A., Kemp B.E.;
 RT "Mammalian AMP-activated protein kinase shares structural and
 functional homology with the catalytic domain of yeast Snf1 protein
 kinase.";
 RL J. Biol. Chem. 269:2361-2364(1994).
 CC -1- FUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT
 INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY
 PROTEIN SNF1. INTERACTS ALSO WITH S1P1, S1P2 AND GAL83. COULD
 PHOSPHORYLATES CAT8.
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC SNF1 SUBFAMILY.
 CC -----
 CC This SNF1-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M13871; AAA35058.1; -;
 CC EMBL: U33050; AAB64904.1; -;
 CC PIR: A26030; A26030.
 CC HSSP: P24941; 1HCL.
 CC GSD: S0002885; SNF1.
 CC InterPro: IPR000719; Euk_Pkinase.
 CC InterPro: IPR002290; Ser_thr_Pkinase.
 CC Pfam: PF00069; Pkinase; 1.
 CC ProDom: PD000001; Euk_Pkinase; 1.
 CC SMART: SM00220; S_TKc; 1.
 CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 CC Transferrase; Serine/threonine-protein kinase: ATP-binding;
 KM Phosphorylation; Cardiolipin metabolism; Nuclear protein.
 KM DOMAIN 18 32 POLY-HIS.
 FT DOMAIN 55 306 PROTEIN KINASE.
 FT NP_BIND 61 69 ATP (BY SIMILARITY).
 FT BINDING 84 84 ATP (BY SIMILARITY).
 FT ACT_SITE 177 177 BY SIMILARITY.
 FT MOD_RES 210 210 PHOSPHORYLATION (AUTO-).
 FT SEQUENCE 633 AA; 72045 MW; F5C633565C986C4E3 CR64;
 Query Match 18.3%; Score 753.5; DB 1; Length 633;
 Best Local Similarity 32.8%; Pred. No. 1.4e-23;
 Matches 166; Conservative 103; Mismatches 144; Indels 93; Gaps 14;

14 AYHLRPHRPPPOHAQ-----YGPYRLKTLGQGTGLKLVHNCI 55
 17 SHHHHHHHHHHHHGGSGNSTLNPKSLADGAHIGNYOIVKTLGSGSGKVLAVHTT 76
 56 TGQKVALIKYVREKLSVLM-KVEREALIKLLENHVLKLDHYENKVKLYLVLEHS 114
 77 TGQKVALIKYVREKLSVLM-KVEREALIKLLENHVLKLDHYENKVKLYLVLEHS 135
 115 GGELFDYLVKGRLLTPREKRFQIVSALDFCHSYSGCHRDLPENLLDEKNNIRAD 174
 136 GNEPFDYLVKGRLLTPREKRFQIVSALDFCHSYSGCHRDLPENLLDEKNNIRAD 195
 175 FGASLQVDSLETSCTGSPHYACPEVYKGEYGRADAMSCGYLLFALLVGLPFPDD 234
 196 FGLSNIMTDGNEFLTKSGSPHYACPEVYKGEYGRADAMSCGYLLFALLVGLPFPDD 255
 235 NLRLDLEKVKRGVHMPFIPDPDQSLRGMEVEPEKLSLEQOKRPTVLGKHE--- 291
 256 SIPVLEKISNGVYTLPELSPGAAGLIRMLVPLARISHEIMODDKYDLPETYL 315
 292 -PDCLPEAPGRVAMR-----SLPSNGELDPVLESMAISGCFRDERLRLRELSSEE 344
 316 PPD-LKHPPEENENNNKSKDGSSPDNDDEIDNLVNLISPMGY-EKDEIYESLESSED 372
 345 ---NQEMITYLLDKREKRPSCEDOLPFRNDY-----DPP-----RK 380
 373 TPAPNEIDAYMLIKENKSLI-----KDMKANKSVSDELDYFLOSPPFQOQSKSHQS 427
 381 RVDSPMLSRHGRKRPERSKMEVLSITDAGGSGSPVPRTRALE---MCHORSRSVSGA 436
 428 QVDHETAKOHARM-----ASATQGRTHYQSPFMQYKREDSVTSIL 470
 437 STGL-----SSPLSPRSV 452
 471 PTLPLQIRANMLAQSPASAKISPL 496
 RESULT 11
 SNF1_CANTR STANDARD; PRT; 619 AA.
 ID SNF1_CANTR 094168;
 AC 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Carbon catabolite derepressing protein kinase (EC 2.7.1.-).
 GN SNF1.
 OS Candida tropicalis (Yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5482;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kanai T., Ogawa K., Ueda M., Tanaka A.;
 RT "Genetic evaluation of the function of SNF1 in Candida tropicalis.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT
 INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY
 PROTEIN SNF4. COULD PHOSPHORYLATES CAT8 (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE (BY
 SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC SNF1 SUBFAMILY.
 CC -----
 CC This SNF1-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB024535; BAA75689.1; -;
 CC HSSP: O63450; IA06.

DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase.1.
 DR ProDom: PD000001; Euk_pkinase.1.
 DR SMART: SM00220; S_TKc.1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST.1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
 DR Transferrase: Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation; Carbohydrate metabolism; Nuclear protein.
 FT DOMAIN 17 29 POLY-HIS.
 FT NP_BIND 52 303 PROTEIN KINASE.
 FT BINDING 58 66 ATP (BY SIMILARITY).
 FT ACT_SITE 81 81 ATP (BY SIMILARITY).
 FT MOD_RES 174 174 BY SIMILARITY.
 FT MOD_RES 207 207 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 SQ SEQUENCE 619 AA: 70323 MW: 0FCF1FC3DCE706D7 CRC64;

Query Match 17.8%; Score 730; DB 1; Length 619;
 Best Local Similarity 37.8%; Pred. No. 1,2e-22;
 Matches 153; Conservative 90; Mismatches 118; Indels 44; Gaps 10;

OY 15 YLPHPHPHPOHAQ-----YVGRLEKTLKRGKGLVLYGVYCTGKQVAVI 62
 DB 21 HHHHHHHHSQPAQPIPIIDPNVNPANRIGRGYQITLIGSGSGKVKLAQVGTGQVAL 80
 OY 63 KLVNREKSESLVM-KVEREIALKLEIHPVYLKLDVYENKLYLYLEHVSQGELEDY 121
 DB 81 KLVNREKSESLVM-KVEREIALKLEIHPVYLKLDVYENKLYLYLEHVSQGELEDY 139
 OY 122 LVKRGRLPRKARKFRQVLSALDFCHSYSGHRLDKPEKLLDEKNNITADPGMSIQ 181
 DB 140 IVORCKMPEDARFEEQOITIAVEYGHRRKIYNRDLKPEKLLDDQNTAVIADPGLSNM 199
 OY 182 VGDSLLESCGSPHACPRVYIGEKYDGRADMSGCVILFALLGALPPDDNLRLQLE 241
 DB 200 TDGNLTKTSCGSPNAAPEVIGSKLYAGPEVYVWSSGVILYVLCRLPPDDFIPALFK 259
 OY 242 KVRKGVFMPHPIPPDCOSILRGMIETVEPKRLSLQIOKHMPYLDGKHEHPPDCLEPAPG 301
 DB 260 KISNGVYLPNPLSGAHNLLRLMVLVNPANRITITHEIMEDWF--KQMPDYLLPPD-- 315
 OY 302 RRVAKRSLPSNGELDPVLESMA-SLGCFFRDR-----ERLARE-----LASEENOE 347
 DB 316 --LSKIKTSIDIDEDVIALSYVMGYDRDEITISYERKANDEAAGAPPTNGSKSTNE 372
 OY 348 KMIYVLLDRKERYPSCHDOLPPRNDV-----PPKRVDSPP 385
 DB 373 VLDAYVLLM--KENHTLVLDLKKSKSENIESITSLSPSSSFPNP 415

RESULT 12
 AAK1_HUMAN STANDARD; PRT: 550 AA.
 AC G13131; 000286; GQGNQ4;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 5'-AMP-activated protein kinase, catalytic alpha-1 chain (EC 2.7.1.-)
 GN (AMPK alpha-1 chain).
 DE PRKAA1 OR AMPK1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Mammary gland;
 RA Yano K.;
 RT "Nucleotide sequence of cDNA for human AMP-activated protein kinase
 alpha-1.";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 [2]

RP SEQUENCE FROM N.A.
 RC TISSUE-Blood;
 RX MEDLINE=20499367; PubMed=11042152;
 RA Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G.,
 RA Shen Y., Pan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W.,
 RA Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.;
 RT "Cloning and functional analysis of cDNAs with open reading frames for
 300 previously undefined genes expressed in CD34+ hematopoietic
 stem/progenitor cells.";
 RL Genome Res. 10:1546-1560(2000).
 [3]
 RP SEQUENCE OF 27-200 FROM N.A.
 RC TISSUE-Intestine;
 RA Taboada E.N., Hickey D.A.;
 RL Submitted (Apr-1995) to the EMBL/GenBank/DBJ databases.
 [4]
 RP SEQUENCE OF 294-550 FROM N.A.
 RC TISSUE-Liver;
 RX MEDLINE=96132781; PubMed=8557660;
 RA Stapleton D., Mitchell H.I., Gao G., Widmer J., Mitchell B.J.,
 RA Teh T., House C.M., Fernandez C.S., Cox T., Witters L.A.,
 RA Kemp B.E.;
 RT "Mammalian AMP-activated protein kinase subfamily.";
 RL J. Biol. Chem. 271:611-614(1996).
 CC -1- FUNCTION: RESPONSIBLE FOR THE REGULATION OF FATTY ACID SYNTHESIS
 BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. IT ALSO REGULATES
 CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND INACTIVATION OF
 HORMONE-SENSITIVE LIPASE AND HYDROXYMETHYLGUTARYL-COA REDUCTASE.
 CC APPEARS TO ACT AS A METABOLIC STRESS-SENSING PROTEIN KINASE
 SWITCHING OFF BIOSYNTHETIC PATHWAYS WHEN CELLULAR ATP LEVELS ARE
 DEPLETED AND WHEN 5'-AMP RISES IN RESPONSE TO FUEL LIMITATION
 CC AND/OR HEPHOKIA. THIS IS A CATALYTIC SUBUNIT (BY SIMILARITY).
 CC -1- SUBUNIT: HETERODIMER OF A CATALYTIC SUBUNIT (BY SIMILARITY).
 CC NON-CATALYTIC SUBUNITS
 CC -1- P1M: AUTOPHOSPHORYLATED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC SNF1 SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: AB022017; BAA36547.1; -
 DR EMBL: AF100763; AAD43027.1; -
 DR EMBL: U22456; AAA64850.1; -
 DR EMBL: Y12856; CAAT3361.1; -
 DR HSSP: O63450; IA06.
 DR GeneW: HGNC:9376; PRKAA1.
 DR MIM: 602739; -
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase.1.
 DR ProDom: PD000001; Euk_pkinase.1.
 DR SMART: SM00220; S_TKc.1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST.1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
 KW Transferrase: Serine/threonine-protein kinase; Fatty acid biosynthesis;
 KW Phosphorylation; ATP-binding.
 FT DOMAIN 18 270 PROTEIN KINASE.
 FT NP_BIND 44 32 ATP (BY SIMILARITY).
 FT BINDING 47 47 ATP (BY SIMILARITY).
 FT ACT_SITE 141 141 BY SIMILARITY.
 FT MOD_RES 174 174 PHOSPHORYLATION (AUTO-) (POTENTIAL).
 FT CONFLICT 133 133 T -> A (IN REF. 3).
 FT CONFLICT 133 133 A -> V (IN REF. 3).
 FT CONFLICT 199 199 I -> L (IN REF. 3).
 FT CONFLICT 260 260 S -> T (IN REF. 2).
 SQ SEQUENCE 550 AA: 62793 MW: 3316183D744DE325 CRC64;

DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 DT Probable serine/threonine-protein kinase YCL24W (EC 2.7.1.1-).
 GN YCL24W OR YCL24W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycos.
 RX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE OF 1-569 FROM N.A.
 RA Duesterhoft A., Erdmann D., Hegemann J., Philippson P.,
 RA Schweitzer B., Spiegelberg R.;
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 567-915 FROM N.A.
 RA Oliver S.G., Anwar R., Brown A., Gent M.E., Indge K.J., James C.M.,
 RA Staveva L.I.;
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP REVISIONS.
 RA Gromada R.;
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC NIM1 SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch.)

 CC EMBL: X59720; CAA24361.1;
 DR PIR: S19351; S19351.
 DR HSSP: Q63450; IAO6.
 DR SGD: S0000529; YCL024W.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR PRODOM: PD000001; Euk_pkinase; 1.
 DR SMART: SM00220; S_TKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
 KW ATP-binding.
 FT DOMAIN 21 285 PROTEIN KINASE.
 FT NP_BIND 27 35 ATP (BY SIMILARITY).
 FT BINDING 50 50 ATP (BY SIMILARITY).
 FT ACT_SITE 152 152 BY SIMILARITY.
 FT SEQUENCE 915 AA; 102688 MW; BFB01C8CA43AC181 CRC64;
 SQ
 Query Match 17.6%; Score 722; DB 1; Length 915;
 Best Local Similarity 38.3%; Pred. No. 3.5e-22;
 Matches 171; Conservative 80; Mismatches 155; Indels 40; Gaps 11;
 QY 31 VGPRLKLTGKGTGLVKGVCITGKVAIKIVNREKLS-----ESVL-MKERE 81
 DB 18 IGPWKLGTETLGFSTGKVOAQHGRHTAVKVISKLFNNNGNSNDSDVLPYNIERE 77
 QY 82 TAILKLHPVLAHLDVYENKKYLYLVHVSGLLEPYLVKKGLTPREARKFRQIV 141
 DB 78 IVIKLLSHPNVLSLYDWETNNNNLYLLEYAEKGEFLMLVDHGPLPREALNCFRQII 137
 QY 142 SALDFCHSYICHRDLDEKNNIRIADFGWASLYQVDSLETSCGSHYACREV 201
 DB 138 IGISYCHALGIYHDLKRENLLDSFYNIKIDFGWALQTPADLETSCGSHYAPPEL 197
 QY 202 IKGERYDGRADWMSGVILFALLVGLAFDD--NLROLKRVKGVNMBH--FIPDP 257

DB 198 VSLGFYEGFASDVMSGVILFALLTGLRPPDEENGVRDILLKQKGFEMPNDTEISR 257
 QY 258 CQSLFGMEVEERKSLSEIQKHRYLGCKHE---PDCLEPAPRRVAMSLP---- 310
 DB 258 AADLLKILVDPDRQRIKINDILISH--LLKKYQITDKSISKIPLPRENTIYLPADSNN 315
 QY 311 -SNGELDPLVLESMASLGCPRDRERLHRESEBNEKMIYIYLLRKERYPSCEODL 369
 DB 316 HTSASIDDSITLQNVYLMHGRHADIVSKLKGNTKELITALLY--RKLDSVRSNK 373
 QY 370 PPRNDVPPRRVDSFMLSNGKRPRERKSMELSTIDAGGSPVPTRALMAHQSR 429
 DB 374 KNRNKKIKTKKRNKRSSTLSSSSSLINNRSIQ-----SRPRRTSK--RHSRE 419
 QY 430 -SRYSVGASTGLSSPSLSPSPVFS 454
 DB 420 FSSSRKRSFLLSNPTDSSPIPLRS 445
 RESULT 15
 SNFL_CANGA
 ID SNFL_CANGA STANDARD; PRT; 611 AA.
 AC 000372;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Carbon catabolite depressing protein kinase (EC 2.7.1.1-).
 GN SNFL
 OS Candida glabrata (Yeast) (Torulopsis glabrata).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 RX NCBI_TaxID=5478;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=NCCL84;
 RC MEDLINE=97101049; PubMed=8945576;
 RA Peltier R., Kwon-Chung K.J.;
 RT "Disruption of the SNF1 gene abolishes trehalose utilization in the
 RT pathogenic yeast Candida glabrata."
 RL Infect. Immun. 64:5269-5273(1996).
 CC -1- FUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT
 CC INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY
 CC PROTEIN SNF4. COULD PHOSPHORYLATES CAT8 (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC SNF1 SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch.)

 CC EMBL: L78130; AAB48642.1;
 DR HSSP: P24941; IHLCL.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR PRODOM: PD000001; Euk_pkinase; 1.
 DR SMART: SM00220; S_TKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation; Carbohydrate metabolism; Nuclear protein;
 FT DOMAIN 6 17
 FT NP_BIND 39 290 PROTEIN KINASE.
 FT BINDING 45 53 ATP (BY SIMILARITY).
 FT ACT_SITE 161 161 ATP (BY SIMILARITY).
 FT ACT_SITE 161 161 BY SIMILARITY.

FT	MOD_RES	SEQUENCE	194	194	PHOSPHORYLATION (AUTO-) (BY SIMILARITY)
50	611	AA; 70049	MM; 89E17812A4900CD0	CRC64;	

SO SEQUENCE - 611 AA; 70049 MW; 89E17812A4900CD0 CRC64;

Query Match	Score	DB 1;	Length
17.58;	720;	DB 1;	611;

Matches 148; Conservative 97; Mismatches 129; Indels 42; Gaps 10;

Matches 148; Conservative 97; Mismatches 129; Indels 42; Gaps 10;

```

QY 15 YLHPRHPRPQNAOY-----VGPRLKTELGKGGTGVLYGVNHTQCA 61
Db 7 HHHHHHHHHHSSNYSNKYSVSLDQSRGNYQYKTLDEGSGFKKLAHYHTQCA 66

QY 62 IKIYNREKLSBSYLM-KVEREIALIKLIEHPHYLKLHDYENKRYLYLVEHSGELED 120
Db 67 LKSLINKVYKASDMQGRIDREISYLRRLRPHILIKLYDIKSDIELIMVIEY-AGNELED 125

QY 121 YLYKKGRLTPREAKREFRQYASLDGFSHSYICHRLDKPKNLLDEKNRIKRIDEFGMASL 180
Db 126 YIYORNMKSEDEARFPOQIIISAEYCHRRKIYHRDLKPEMLLDEHLANKIADFGLSMT 185

QY 181 QVGSILLETSGSHYACPEYIKGEKEDGRRADMWSCVTLFALLVGALEPDDDNIROLL 240
Db 186 MTDGNSFLKTCGSGSNYAPAYISKLYAGCEVDWWSGVTLIYMLCRRLPFDEESTIYLF 245

QY 241 EKVARGYEHMHFIPRPCQSLNRCMIEVEEPKRLSLBQIKHPMYLGCKHEPDPCEPAP 300
Db 246 KNISNGVYTLPLKFLSPGASDLIKMLLVNPLNRISIHENQDEMF---KYDLAEVLYPDQ 302

QY 301 GRVYAMSLPSNG-----ELDDPYLESMA-SLGCFRDRERLHRLHSEEE---NDEKMT 350
Db 303 LKOOEOPFKKSGNEENVEIDEXVYVLSKTMG--YDKDEIYALBESSEDTPAYNEIRMA 360

QY 351 YLLILDRKERAPSCED-----QDL-----PPANDVDPKRRVDSFMLSRHGKRR 394
Db 361 YLLIDNKSILIKMDKNNTYQEDLTLSOSPFPFOONGGAKASEQOKKHHSGRR 416

```

Db 7 HHHHHHHHNSGSYVSNKVSSLDGSRVGNQIVKTIGEGSF GKVKLAYHVTGGQVA 66

62 IKIVNREKLSVLM-KVEREIAIKLIEHPVLKLDVYENKKYLYLVEHVS GGELEFD 1200

Db 67 LKSINKKVLAKSDMGRIDREISYLRLLRHPHIKLYDVIKSKDEIIMVIEY-AGNELFD 125

```
OY      121  XLVKKGRLTPKEARKEFRQIVSALDFCHSYSICHRDLKPENLLDEKNRIRIADFGMASL 180
```

Db 126 YIVQRNKMSEQEARRFQQIISAVEYCHRHKLVHRDLKPENLLDLDEHLNVKIADFGLSNI 183

181 QVGDSLEISGSPHIACFEYINGENIDGRKADMMSCVILFALLVGALEFDDNNKLL 24

DD 100 MLDGNE LK15CGSFNIAFEV13GNLHGFEEVAVMSGGV LTI VWECKNNE F EDECELE VLE 27

[illegible][illegible][illegible][illegible]

```

361 VTTRDKRSITKDMKODNNVTQELDTETSCSPDTEONGDGMKASEDOKKHSGRB 416

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Search completed: April 16, 2003, 12:30:21
Job time : 34 secs

Job time : 34 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2003, 12:21:57 ; Search time 43 Seconds
(without alignments)
2410.907 Million cell updates/sec

Title: US-10-003-690-2

Perfect score: 4108
Sequence: 1 MSGGAKGGGGSPAHPLPH.....PRGGPKKLLATNGTPLP 778

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 segs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : A_Geneseq_101002.*

1: /SID2/gcgdata/geneseq/genesep-emb1/AA1980.DAT.*
2: /SID2/gcgdata/geneseq/genesep-emb1/AA1981.DAT.*
3: /SID2/gcgdata/geneseq/genesep-emb1/AA1982.DAT.*
4: /SID2/gcgdata/geneseq/genesep-emb1/AA1983.DAT.*
5: /SID2/gcgdata/geneseq/genesep-emb1/AA1984.DAT.*
6: /SID2/gcgdata/geneseq/genesep-emb1/AA1985.DAT.*
7: /SID2/gcgdata/geneseq/genesep-emb1/AA1986.DAT.*
8: /SID2/gcgdata/geneseq/genesep-emb1/AA1987.DAT.*
9: /SID2/gcgdata/geneseq/genesep-emb1/AA1988.DAT.*
10: /SID2/gcgdata/geneseq/genesep-emb1/AA1989.DAT.*
11: /SID2/gcgdata/geneseq/genesep-emb1/AA1990.DAT.*
12: /SID2/gcgdata/geneseq/genesep-emb1/AA1991.DAT.*
13: /SID2/gcgdata/geneseq/genesep-emb1/AA1992.DAT.*
14: /SID2/gcgdata/geneseq/genesep-emb1/AA1993.DAT.*
15: /SID2/gcgdata/geneseq/genesep-emb1/AA1994.DAT.*
16: /SID2/gcgdata/geneseq/genesep-emb1/AA1995.DAT.*
17: /SID2/gcgdata/geneseq/genesep-emb1/AA1996.DAT.*
18: /SID2/gcgdata/geneseq/genesep-emb1/AA1997.DAT.*
19: /SID2/gcgdata/geneseq/genesep-emb1/AA1998.DAT.*
20: /SID2/gcgdata/geneseq/genesep-emb1/AA1999.DAT.*
21: /SID2/gcgdata/geneseq/genesep-emb1/AA2000.DAT.*
22: /SID2/gcgdata/geneseq/genesep-emb1/AA2001.DAT.*
23: /SID2/gcgdata/geneseq/genesep-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	3956.5	96.3	794	22	AAU03517
2	3956.5	96.3	794	23	AAE16271
3	2794	68.0	674	23	AAAM7830
4	2767.5	67.4	664	23	AAE21723
5	1952	47.5	851	22	ABE62061
6	1341.5	32.7	301	22	ABBI1052
7	1043	25.4	204	22	ABBI1680
8	862	21.0	1349	22	AAE70854
9	851	20.7	1369	23	AAE21712
10	840.5	20.5	823	21	AAW90879

11	831.5	20.2	783	22	AAE65764	Human protein kina
12	831.5	20.2	783	22	AAE85786	Human kinase PKIN-
13	828.5	20.2	768	22	AAW78349	Human protein SEQ
14	828	20.2	691	23	AAE19050	Human PAR-1B alpha
15	828	20.2	724	23	AAE19051	Human PAR-1B beta
16	826	20.1	1251	22	AAE65626	Novel protein kina
17	824.5	20.1	790	21	AAW90878	Human keratinocyte
18	823	20.0	688	23	ABE04434	Human neuronal ser
19	823	20.0	752	22	AAE11782	Human kinase (PKIN
20	823	20.0	752	22	AAW93956	Human polypeptide,
21	823	20.0	752	23	ABE04433	Human neuronal ser
22	822	20.0	804	22	AAW79333	Human protein SEQ
23	820	20.0	722	17	AAW98227	Rat neuronal prote
24	816	19.9	786	22	AAU03518	Human protein kina
25	806.5	19.6	795	23	AAE19052	Human PAR-1C prote
26	806.5	19.6	795	23	AAE16258	Human kinase PKIN-
27	803.5	19.6	660	23	ABE04432	Murine neuronal se
28	801	19.5	724	23	ABE04431	Murine neuronal se
29	795	19.4	729	19	AAW37158	Human Twenty-five
30	795	19.4	729	22	AAE65628	Human protein kina
31	795	19.4	782	22	AAW47219	Human NOVI0 protei
32	792	19.3	744	23	AAE19049	Human PAR-1A prote
33	789	19.2	514	21	AAE03420	Soybean putative c
34	789	19.2	926	22	AAE65631	Novel protein kina
35	786	19.1	515	21	AAE03422	Soybean putative c
36	786	19.1	1102	21	AAE43286	Human ORF3050
37	776	18.9	520	22	AAE65629	Novel protein kina
38	767	18.7	1398	22	ABE60535	Drosophila melanog
39	766.5	18.7	511	14	AAE40842	SHP. Yeast sp.
40	764	18.6	739	22	ABE63486	Drosophila melanog
41	763.5	18.6	779	17	AAW98226	Rat neuronal prote
42	762.5	18.6	832	22	AAE68340	A human dishevelle
43	762.5	18.5	832	23	AAE19053	Drosophila sp. PAR
44	761.5	18.5	512	21	AAE6157	Arabidopsis thalia
45	761.5	18.5	512	23	AAE17663	A thaliana AKIN11.

ALIGNMENTS

RESULT 1	AAU03517	standard; Protein: 794 AA.
ID	AAU03517	
AC	AAU03517	
XX		
DT	12-SEP-2001 (first entry)	
DE	Human protein kinase #17.	
XX		
KW	Human; protein kinase; PTK; STK; cancer; cardiovascular disease;	
KW	metabolic disorder; immune related disease; neurological disorder;	
KW	neurodegenerative disorder; inflammatory disorder; infectious disease;	
KW	reproductive disorder.	
XX		
OS	Homo sapiens.	
XX		
PN	MO200138503-A2.	
XX		
PD	31-MAY-2001.	
XX		
PF	22-NOV-2000; 2000MO-0532085.	
XX		
PR	24-NOV-1999; 99US-0167482.	
XX		
PA	(SUGR-) SUGEN INC.	
XX		
PI	Flowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;	
PI	Flanagan P, Clary D;	
XX		
DR	WPI: 2001-343950/36.	
XX	N-PSDB; AAS06717.	
XX		


```

Db 71 HPHVTKLADHYENKYYLYLVLEHVSQGLLFDYLVKKGRLLTPKARFFROIISALDFCHS 130
Oy 150 YSICHRDLKPEMLLDEKNNIRIADFGMASLQVDSLLETSCGSPHYACPEVIGKEKYD 209
Db 131 HSHCHDLKPEMLLDEKNNIRIADFGMASLQVDSLLETSCGSPHYACPEVIGKEKYD 190
Oy 210 RRADMVSCGVLLEFALLVGLPDDNLRLLEKVKRQVFMHPIPPDCOSLLRGMITE 269
Db 191 RRADMVSCGVLLEFALLVGLPDDNLRLLEKVKRQVFMHPIPPDCOSLLRGMITE 250
Oy 270 PERKRLSDIOKHPWLVGKKEHPDCLERAPGRVAMRSIPNGELDDPDVLSMASIGCF 329
Db 251 AARLLELHIOKHWYIGGNEPEP-EQPIP-RKVOIRSLLEDIDPDVLSMASIGCF 308
Oy 330 RDERLARELSEENQEKMIYLLDRKERYPSCEDQDPPRNDVDPKRRVDSPLMSR 389
Db 309 RDRNKLQDLSSEENQEKMIYLLDRKERYPSCEDQDPPRNEIDPPKRRVDSPLMSR 368
Oy 390 HGRRPERRKSMVLTSTDAAGGSPVPTRRALMAOHQSRSVSGASTGLSSPLSSPR 449
Db 369 HGRRPERRKSMVLTSTDAAGGSPVPTRRALMAOHQSRSVSGASTGLSSPLSSPR 425
Oy 450 SPFESRSPERGAGDEARGGSPSTKQTLPSRGRPGGAGEOPPPPARSTPLPGPGSP 509
Db 426 -----VTPHSPRSGSPLETPKRG-- 442
Oy 510 RSSGCTPLHSPHTPRAPPTGPTGTPPPSPGGGAGVGAAMRSRLNSIRNGLSPRRHR 569
Db 443 -----TPVHTPKRSPAGTNPPTPSPSP--SVGGVPPRARLNSIKNSPFLSGPRHR 491
Oy 570 KMOVPPAEKMSITPSSPPLAKRSFNGNISLDEKQITLVLDKPLSLIKADIVAF 629
Db 492 KLOVPPAEKMSITPSSPPLAKRSFNGNISLDEKQITLVLDKPLSLIKADIVAF 551
Oy 630 STPLSHSVYLSOTSPFAREYASGSPVFOKPVRRQVYDSSSEGEPPRRDSSGGGYS 689
Db 552 STPLSHSVYLSOTSPFAREYASGSPVFOKPVRRQVYDSSSEGEPPRRDSSGGGYS 605
Oy 690 VPEFLISGPRRRKRVETIOAQLLSTHDOPSVYALDEKN 730
Db 606 VPEFLISGPRRRKRVETIOAQLLSTHDOPSVYALDEKN 646

RESULT 5
ABB62061
ID ABB62061 standard; Protein; 851 AA.
XX
AC ABB62061:
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 12975.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
  pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN MO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001MO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
PA (PEKE ) PE CORP NY.
PI Venter JC, Adams M, Li PWD, Myers EW,
XX WPI; 2001-656860/75.
XX DR N-PSDB; ABL06164.
  
```

```

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
PS Disclosure; SEQ ID NO 12975; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB116177), expressed DNA
CC sequences (AB116173-AB116174) and the encoded proteins
CC (AB116173-AB116174).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 851 AA:
XX
Oy 26 OHAYVGRLEKTLGKQGLVGLVHCITGQVAKIYVREKLSVLMKVEREAIL 85
Db 10 ENCQFVGRLEKTLGKQGLVGLVHCITGQVAKIYVREKLSVLMKVEREAIL 69
Oy 86 KLIEHPYKLDHYENKYYLYLVLEHVSQGLLFDYLVKKGRLLTPKARFFROIISALD 145
Db 70 KLIEHPYKLDHYENKYYLYLVLEHVSQGLLFDYLVKKGRLLTPKARFFROIISALD 129
Oy 146 FCHSISICHRDLKPEMLLDEKNNIRIADFGMASLQVDSLLETSCGSPHYACPEVIGKE 205
Db 130 FCHSISICHRDLKPEMLLDEKNNIRIADFGMASLQVDSLLETSCGSPHYACPEVIGKE 189
Oy 206 KYDGRADMVSCGVLLEFALLVGLPDDNLRLLEKVKRQVFMHPIPPDCOSLLRGMITE 265
Db 190 KYDGRADMVSCGVLLEFALLVGLPDDNLRLLEKVKRQVFMHPIPPDCOSLLRGMITE 249
Oy 266 IEVEPERKLSLEOIKHWY-LGGKHEPD--PCLEPARGRVAMRSIPNGELDDPDVLE 321
Db 250 IEVEPERKLSLEOIKHWY-LGGKHEPD--PCLEPARGRVAMRSIPNGELDDPDVLE 304
Oy 322 SMASIGCFRDERLARELSEENQEKMIYLLDRKERYPSCEDQD--LPPRND--V 375
Db 305 AICSLGCFRDERLARELSEENQEKMIYLLDRKERYPSCEDQD--LPPRND--V 364
Oy 376 DPPKRRVDSPLMSRHRGRKRRPERSKSEVLTSTDAAGGSPV-PTRRALMAOHS---QRS 430
Db 365 DPPKRRVDSPLMSRHRGRKRRPERSKSEVLTSTDAAGGSPV-PTRRALMAOHS---QRS 413
Oy 431 RSVASGASTGLSSPLSSP--RSVPFS-----FSPPEGA-----GDEARG--GS 470
Db 414 RSVASGASTGLSSPLSSP--RSVPFS-----FSPPEGA-----GDEARG--GS 473
Oy 471 PTKTQTLPS-----RGRGGAGAGQPPPARSTPLPG-PPG 507
Db 474 SVSRTPSHSQSLIEGDVYVYVREPRTERDRLRGRGG-----SPRBDGDIPIPG 525
Oy 508 SP--RSSGCTPLHSPH-----TPRASP--TGTPGTTPPSPGGAGGAM 549
Db 526 SPGNSGSGTSPASPSVHRANSOPTIATVNPNGSPMANNSPGM--PGSPONTPGOLW 583
Oy 550 RSLNSINSFSGSPRRRRKQVPTAEEMSLTPESSPELAKRSFNGNISLDEKQIT 609
Db 584 KTRLNINNSFSGSPRRRRKQV-NADEV-HLTPESSPELAKRSFNGNISLDEKQIT 641
Oy 610 LVLDKPLSLIKADIVAFPLSLSHSVLSOTSPFAREYASG-GPSVFOKPVRRQVDS 668
Db 642 ILVKGKPIATYAKHLIHAFLSMALSHSVLSOTSPFAREYASG-GPSVFOKPVRRQVDS 701
  
```

OY 669 SSGPPSPRRDGGGGGIVSVTTLISGPRRRKRVETIOAQLLS 715
 DB 702 AI-----CKGDIDMFLPFLTFITLISGNIRRRRCHEHQVCS 741

RESULT 6

ABBI1052
 ID ABBI1052 standard; peptide; 301 AA.

XX ABBI1052;

DT 11-JAN-2002 (first entry)

DE Human HRPDK-1 homologue, SEQ ID NO:1422.

XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritis; haemostatic; antiarteriosclerotic;
 KW cytoskeletal; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antitumor.

OS Homo sapiens.

XX WO200157188-A2.

PD 09-AUG-2001.

PF 05-FEB-2001; 2001WO-US03800.

XX 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

PI (HXSE-) HXSEQ INC.

XX Tang YT, Liu C, Dermanac RT;

XX WPI: 2001-457740/49.

XX N-PSDB; ABA08296.

XX Human proteins and DNA encoding sequences useful for preventing,

XX treating or ameliorating a medical condition in a mammalian subject

XX e.g. arthritis and cancer.

XX Claim 20; Page 140; 1963pp; English.

XX Sequences ABBI0981-ABBI2330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include

CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.

XX Sequence 301 AA;

XX Query Match 32.7%; Score 1341.5; DB 22; Length 301;

XX Best Local Similarity 87.2%; Pred. No. 1.5e-82;

XX Matches 259; Conservative 1; Mismatches 2; Indels 35; Gaps 1;

OY 27 HAQVYPRLEKTLKGGTGLVAGVHCITGOKVAIKIYRKLSSVLMKVRRTAILK 86

DB 1 HAQVYPRLEKTLKGGTGLVAGVHCITGOKVAIKIYRKLSSVLMKVRRTAILK 60

OY 87 LIEHRYVILKHDYENKYL-----YLVLE 111

DB 61 LIEHRYVILKHDYENKYL-----YLVLE 120

OY 112 HVSGETLDYLVKRGRLPPEKARPPROIVSALDPCHSYSTICHRDLKPEMLIDKNNIR 171

DB 121 HVSGETLDYLVKRGRLPPEKARPPROIVSALDPCHSYSTICHRDLKPEMLIDKNNIR 180

OY 172 IADFGMASLQVDSLSLETSCGSPHYACPEVIRGEKYGRRADMSCGVILFALLVGLP 231

DB 181 IADFGMASLQVDSLSLETSCGSPHYACPEVIRGEKYGRRADMSCGVILFALLVGLP 240

OY 232 DDDNRLQLEKVRKGVFMHPIPPDCGSLNGMFEVEPEKRLSLQIOKHPYTLG 288

DB 241 DDDNRLQLEKVRKGVFMHPIPPDCGSLNGMFEVEPEKRLSLQIOKHPYTLG 297

RESULT 7

XX ABBI1680 standard; peptide; 204 AA.

XX AC ABBI1680;

DT 11-JAN-2002 (first entry)

DE Human Ser/Thr kinase homologue, SEQ ID NO:2050.

XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritis; haemostatic; antiarteriosclerotic;
 KW cytoskeletal; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antitumor.

XX Homo sapiens.

XX WO200157188-A2.


```

Db 182 LBSNGKLETTSCGSPHYAPAEIVSGLEKYGASDVMSCGYLTFALLTGRLPEDEENINL 241
OY 240 LERKRGVEHMP-HFTPPDCOSILRGMEVEPEKRLLEOKHPW-----YLGG 288
Db 242 LKLVQGNFEMPVDEYSREARDLIARLEVDPMRISTEKILRHPILTKYPMSEDLISE 301
OY 289 KHEPDCLEPARBARRAMSLPSNGELDPVLSMASLGCPRDRERLHRELSSENOEK 348
Db 302 KSLPHPE-----HTGYSIGSVNRNDKQILNLTLIAMDREPEEIVDCLJGKSNEK 353
OY 349 MIYLLDKEKERYPCGEDDLPFRNDVDPKRYVDSMLSRHGKR-----RPEKSEVYL 403
Db 354 TTYALLM-----KKHMDNN-----TNNSPKKSISFNKKVYRSGSKYILNGTPRRKRASHI 406
OY 404 ST-----TDAGGGGSPVPTRRALEMAQHSQRKSVSGASTGLSSPLSP-RSPV- 452
Db 407 SVSRPTSFQKSNPGAGA-----TANRNSVARHSVASSANNSPKSPYRSYRSYRSPYK 462
OY 453 -----FSESPREGADGAEAGGGS-----PRTSKQTL 478
Db 463 SPSKRYSYNQSPKSPYGRRSNSQROFENEPLAKPRNIYNEIYDAQSNFSLPPLPSL 522
OY 479 PSRGRGAGGAGOPPPSARSTRPLPGPSPRSSGGPLHSPHTPASPCTGTCTPP 538
Db 523 PSKDSR--YMDPEPNQPOLQAPALSOVPEP-----IYDESPDL 559
OY 539 SPGGVGGAAMSRSLNSIRNSFLSGSPFRHRR--MOVPAEEMSLTPSPSELAKRW 595
Db 560 MOSAKISSGRKRSIIIGKNNNSNSKRMKRSIRASMTGLKRNSTMLLTSTYATLNG 619
OY 596 FGNFSLKREQ-----IFLVLDKPLSSIKADIVHAFSLPSLSHS-----VLSQTSF 644
Db 620 DDDWYMDQIKRTSATFALCDKIFNOEDYDEDEQLVDEPEKEKEYERLMELERKKH 679
OY 645 RAERYKASGSPYFQKPVPRQV--DISSESGEPSPRRDGGGGGIVSVFTLLSGPSRRF 702
Db 680 EAELEKAR--RELEKRRKRRKRSILSSKRLSTIVYNDADPNNSGEQVDEGIKOPRKOS 736
OY 703 KRYVETIAQLISTHDQSVQALADEKKAQTPPACAPPSRILOPPGRRDPPELSSSP 759
Db 737 KNLV---ALRALSEGNNHASEELTLEVENIMKRRSASQ-----VPRKRQPTVLTTRP 785

```

RESULT 9
AAE21712
ID AAE21712 standard; Protein; 1369 AA.

XX AAE21712;

DT 16-JUL-2002 (first entry)

DE Human PKIN-7 protein.

XX Human; kinase; enzyme; PKIN-7 protein; immune system disorder; anaemia;
XX acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease;
XX asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease;
XX AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma;
XX leukaemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder;
XX Down's syndrome; gene therapy; protein therapy; cyostatic.

OS Homo sapiens.

XX Key location/Qualifiers

FT Peptide 1..14 /label- signal_peptide

FT Protein 15..1369 /label- Human_mature_PKIN-7_protein

FT Domain 66..317 /note- "Eukaryotic protein kinase domain"

FT Domain 67..308 /note- "Protein kinase domain"

FT Domain 70..305

```

FT /note- "Protein kinase domain"
PN W0200218557-A2.
PD 07-MAR-2002.
XX 31-AUG-2001; 2001WO-US27219.
XX 31-AUG-2000; 2000US-229873P.
XX 08-SEP-2000; 2000US-231357P.
XX 14-SEP-2000; 2000US-232554P.
XX 22-SEP-2000; 2000US-234902P.
XX 29-SEP-2000; 2000US-236499P.
XX 06-OCT-2000; 2000US-238389P.
XX 13-OCT-2000; 2000US-240542P.
XX (INCYTE GENOMICS INC.
XX Bhandman O, Nguyen DB, Walla NK, Hafalia AJA, Yao MG, Gandhi AR;
XX Gururajan R, Ding L, Patterson C, Yue H, Baughn MR, Tribouley CM;
XX Thornton M, Elliott VS, Lu Y, Ison CH, Au-Young J, Tang YT;
XX Azimzal Y, Burdill JD, Marcus GA, Zingler KA, Lu DAM, Lai PG;
XX Rankumar J, Warren BA, Kearney L, Policky JL, Thangavelu K;
XX Butford N;
XX WPI; 2002-329769/36.
XX N-PSDB; AAD34304.
XX New human kinases, useful for diagnosing, treating or preventing immune
XX system disorders (e.g. Crohn's disease), neurological disorders (e.g.
XX leukemia or lymphoma)
XX Claim 62; Page 152-156; 218pp; English.
XX
XX The present invention relates to human kinases (PKIN) and polynucleotides
XX encoding such proteins. PKIN sequences of the invention are useful for
XX diagnosing, treating or preventing disorders associated with aberrant
XX expression of PKIN, particularly immune system disorders (e.g. acquired
XX immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,
XX anaemia, asthma), neurological disorders (e.g. epilepsy, Charcot-Marie-
XX Tooth disease or seizures), cell proliferative disorders (e.g. cancers
XX such as adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma, sarcoma),
XX and developmental disorders (e.g. Down's syndrome). They are also used
XX in gene therapy and protein therapy. The present sequence is human
XX PKIN-7 protein.
XX
XX Sequence 1369 AA:
SQ
Query Match 20.7%; Score 851; DB 23; Length 1369;
Best Local Similarity 29.4%; Pred. NO. 1.2e-48;
Matches 258; Conservative 124; Mismatches 291; Indels 204; Gaps 30;
OY 2 SSGAKEGGGSPAHLPHP-----HHPHQNAQ-----YGPRLKLT 39
Db 12 AAGGTGAGGAGRLLPAPGSPAPAPSPAPQPPAPASGPPAPARGYEIDT 71
OY 40 LKQGTGLVKGVCITGQKVAIKIVNREKLSSEYIMKVERITALIKLEPHVLIKH 99
Db 192 ENLLDANLNKIKADSPSNIETPGQLTKTWCSPPYAPAEIPEGEKEYDGPVDTMSLGV 251
OY 72 IKGKPNFVVKRATHLVYAKAIAKIIDTQDDENLKIFREVOIMKMLCHPHTIRLYOV 131
OY 100 YENKTYLYLVIEHVSGBLEFDYLVKGRLPKREARFRQIYVSLDFCHSYSTRDLKP 159
Db 132 METERMITYLVTEVASGEIIFDHLVAHGMAKEARFKQIYVYFCHCRNIVARDKA 191
OY 160 ENLLDKNRIRIDFGASIQVDSLSLETSCGSHYACPEYIKGEKDGRRADMSGCV 219
Db 192 ENLLDANLNKIKADSPSNIETPGQLTKTWCSPPYAPAEIPEGEKEYDGPVDTMSLGV 251
OY 220 ILFALLVGALEPDDNLRQLEKVRGVFHPHFPPDCOSILRGMEVEPEKRLSTDOI 279
Db 252 VLVYLVGALPFDGSTIQLNLRARVLSGKFRIFPMSTECHEILIRMLVLDPKRLSMQI 311

```

```

QY 280 QKHPTVLGKHEHPDCLERAPGRVA-MRSLPSNGELDP---DYLESMAISGCFDRERL 335
DB 312 CQKRWKMLG---DADPNMD---RLIACQQLKEKRDYDPLNEDVLAMEDMGL--DKBOT 363
QY 336 HRLRSSEENQEKMIYVLLDRKERYSCEDQDP--PR-----NDV 375
DB 364 LQSLRSDAYHYSATLSLDCDRKHKHTLRGALPSMPRALAFQAPVIOAEQAGTAMNI 423
QY 376 DPKRKRVDSFMLSNGKRRERKSM---VLSI--TDAGGGSPVPTRRALMAQHS--- 427
DB 424 SVQVQVLIIN-----DENQIVEPDGTLNLDSDGEESPEALVYLSMRRTVGVA 473
QY 428 -QSRSVSGASTGLSSPSLSPRSVFSFSPERAGDEARGGSTKTQTLPSGPRGG 486
DB 474 DPKTEWEDLQKLLPFGPVNPQAPFLQVAP-----NVNFMNHLPMQNLPT 521
QY 487 GAGEOP-----PPSARSTPLPGPGSPRSRGGT--PLHSP--LHTPRA-SPTGT----- 531
DB 522 GQLEKESQLQPTTDLNMGKPLGRASDGANIQHQAOLKRRPGSPFLVMTAV 581
QY 532 PGTPP--PPSPGGVGAAMRSRLNSTRNFTLGSPTRRKRRMOVPTAEKMSLTPESSP 588
DB 582 PATVPEDESSDGPDPDEAVORYLANSKR-----HTLAMTPRAE---IPDLQR 629
QY 589 ELAK-----KSWGNFTSLKEQIFLYLADKPLSSIK-ADIVH---AFSLPSLSHV 638
DB 630 QLGQOPFRSRVMPHVL-----PDQHRSTYKDSNTLHLPTERFSVRRFSQDA 677
QY 639 LSGTSPFAEKASGGPS-----VFQKPVRFQVDISSSEGEPEPRRDGSGGGIY 688
DB 678 ASIQARAHLEKKNSSIKQLQCEQLOKMYGGID----- 715
QY 689 SVFTLLISGSRREKRVETIQALLSTHDPVQVALADENKGAQTRPAGAP----- 741
DB 716 -----ERTLEKYOQHMLYQEQHQQIIOQOIQDSICPPQSPPLQACEN 761
QY 742 -----RSLQ-----PPGRDPPELSSPRKGP 764
DB 762 QPALTRHQIQLRIQPSPPNHPNHLFRQPSNSP 798

RESULT 10
AAM90879
ID AAM90879 standard; Protein: 823 AA.
XX
AC AAM90879;
XX
DT 07-JUL-2000 (first entry)
XX
DE Human keratinocyte derived pke#122 protein #2.
XX
KW Keratinocyte; regulatory protein; human; pke#122; antiproliferative;
KW kinase; signal transduction; desmosome; dermatological condition;
KW pemphigus vulgaris; psoriasis vulgaris; epidermis; cosmetic.
XX
OS Homo sapiens.
XX
PN MO200017232-A2.
XX
PD 30-MAR-2000.
XX
PE 06-SEP-1999; 99WO-DE02865.
XX
PR 19-SEP-1998; 98DE-1042863.
XX
PA (KRAM/) KRAMER M.
XX
PI Kramer M, Bechtel M, Reinartz J, Schaefer B, Wallich R;
XX
DR WPI; 2000-283542/24.
XX
DR N-PSDB; AAX82952.
XX
PT New regulatory polypeptide from keratinocytes, useful for diagnosis and

```

```

PT treatment of dermatological disorders, also in cosmetics
XX Claim 1; Page 53; 55pp; German.
PS
XX
CC This invention describes a novel human regulatory polypeptide designated
CC pke#122 (I), or its variants from human keratinocytes, which in activated
CC keratinocytes is the same as, or similar to, the most highly expressed
CC protein and has antiproliferative activity. (I) is a kinase involved in
CC signal transduction and is conjectured to be involved in cell-cell or
CC cell-matrix interactions and/or formation of (heml)desmosomes. Modulation
CC of (I) will affect proliferation and differentiation of keratinocytes.
CC (I) is upregulated in some dermatological conditions, e.g. pemphigus
CC vulgaris and psoriasis vulgaris. (I) is used to raise specific antibodies
CC (Ab) which are used for diagnosis and treatment of dermatological
CC disorders and for cosmetic treatment of the epidermis. Nucleic acid (II)
CC that encodes (I) is used; (I) to produce transgenic mammals (particularly
CC mice and rats); and (II) as source of sense and antisense
CC oligonucleotides for diagnosis and treatment of dermatological disorders
CC and for cosmetic treatment of the epidermis. (II), (I) and polypeptides
CC (Ia) produced by splice variants of (II) are used to identify specific
CC binding agents (potentially useful in medicine and cosmetics) that
CC modulate their function and/or expression, especially inhibitors and
CC activators. Cells and animals transformed with (II) are used to study
CC (I)-induced changes in cell morphology and basic functions. This
CC sequence represents the human keratinocyte protein pke#122 described in
CC the method of the invention.
XX
SQ Sequence 823 AA;
XX
Query Match 20.5%; Score 840.5; DB 21; Length 823;
Best Local Similarity 29.9%; Pred. No. 3.2e-48;
Matches 256; Conservative 102; Mismatches 238; Indels 201; Gaps 27;
QY 4 GAKGCGGSAVYHL-----PAPHPPQHAQVYGRYRLEKTLGKGTGLVKGVC 55
DB 29 GRSGGGGAGAMVMEFASADPAGQSGQKSLRVGYDLERTLGKPNFVYKLARHY 88
QY 56 TGQKVAIKIYNREKLSSVLMKVERETAIKLEIHPYVLKLDHYENKKYLYLVHVS 115
DB 89 TKTOVAIKIITDKTRLDSNNLEKIYREVOLMKLHNPILIKIYQWETMDMYIYVEFAKN 148
QY 116 GELFDVLYKGRLLPKERKREFRQIVASLDFCHSYSGHGRDLKPNLLDEKNRIRDI 175
DB 149 GEMFDYLTSGHLSSENDERKFKFQILSVAYECHDHVYHRLKTNLLDGMKIDLF 208
QY 176 GMAISQVDSLLETSCGSPHYACEVYKGEYDGRADWMSGCVLLFALLVGALEPDDN 235
DB 209 GFGNFTYSGELSTWCGSPYAAPEVFEKGEYEQDILWSLGYLVLYVCGSLPFDGN 268
QY 236 LRQLEKVKRGVFMPIFIPDCOSLRGMTEVEPEKLSLEQLOKHWYLAGKEPPDC 295
DB 269 LPTLRQVYLEGRFRIFPFMSQDCESLIRMLVVPARITTAQINQHRMM-----RAEPC 323
QY 296 LEPAPG-RVAMRSLPSN-GEIDPDVLESMAISGCFDRERLHRLRSSEENQEKMIYV 353
DB 324 L-PEPACPAFSAHYTSNLSGDIYDEQALGIMOTLGY--DRQVYESLQNSYHFAIYTL 380
QY 354 LLDKERYPSCEDQDLPRLNDVDPKRRKRVSPMLSRHG-KRRPERKSMELVSLTDAGGG 412
DB 381 LERLKEKRYNQ-----CARPEPARQPRRSSDLSLEVQEG 419
QY 413 SPVPTRRALMAQHSQNSKSVSGASTG--LSSS-----PLSSRSPVFSFSP----- 457
DB 420 STDPRPALLCPOPOTLVQSVLAQEMDELSLQWLPFFPDADACSGFRPVPSSSL 479
QY 458 -EPGAGDEARGGSPSTKTOTLPSNGPRGGGAGEQPPPSASRTPLPGP----- 505
DB 480 LDRAISEPARQPGLEEQDQTESL-PSSTGRHRTLAESTRLSLTAPCIYVSTTAS 538
QY 506 -----PGSPRSSGGTPLH-----SPLHTPRASP-TGTPGTTPPPSPGG 542
DB 539 PABGTSSDSCLTFSASKSPAGISGTPTATGGLGACSPVRL--ASPLFGQSATPVLAQGG 596

```

QY 543 GVGGA-----AMSRRLNSIRNS--FLGSPRRH---RRKMOV 574
 DB 597 GLGGAVALPVSEFGRRASDTSLTGLKAFQQLRKTTRTGFLGLNKIKGLARQCV 656
 QY 575 TAEEMSS-----LTPESSPEL-----AKRSMFGNFISLDKEQFVLKDKPLSIKAD 623
 DB 657 ASRASRGSLSPFHAAPSPGLHGAAGSREGW----- 688
 QY 624 IYHAFSLPISLASHSVLSQTS- FRAEYKASGGSPVQKPRFOVDISSEGEPPRR--- 679
 DB 689 -----SLEEVLEEQRLQLQHHPAAAP-----GCSQAPQAPAPAPVY 726
 QY 680 ---DSGGGGIVSVTF---LISGPRRRFRVETIOAQLSTHDPQSVQALADEKNGA 732
 DB 727 APCDGFGAAPLPSTLLTSGLPPLPPLQLGTGASPVASAAQLDTH-----LHIGTG 778
 QY 733 QTRPAGAPRRSLQPPG 749
 DB 779 TALPAVPPPRRLARLAPG 795

RESULT 11
 AAG65764
 ID AAG65764 standard; Protein; 783 AA.
 AC AAG65764;
 DT 07-JAN-2002 (first entry)
 DE Human protein kinase polypeptide 3714.
 XX
 KM cytosolic; antiinflammatory; immunosuppressive; cardiant; hepatotrophic;
 KW gene therapy.
 OS Homo sapiens.
 PN MO200173050-A2.
 PD 04-OCT-2001.
 PF 23-MAR-2001; 2001WO-US09483.
 PR 24-MAR-2000; 2000US-191846P.
 PA (MILL-) MILLENNIUM PHARM INC.
 PI Meyers R;
 DR WPI; 2001-611632/70.
 DR N-PSDB; AAI66822, AAI66823.
 PT New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
 PT useful in diagnosis of cancer or cellular proliferation or
 PT differentiation disorders and to screen for polypeptide modulators
 PT useful to treat such conditions
 PS Claim 4; Fig 1A-D; 169pp; English.
 XX
 CC The invention provides novel human protein kinase polypeptides, 3714,
 CC 16742, 23546 and 13887 and nucleic acid molecules encoding them. The
 CC protein kinase polypeptides can be expressed by standard recombinant
 CC methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides
 CC are useful for diagnostic and screening methods to identify subjects (at
 CC risk of) having cancer or cellular proliferation and/or differentiation
 CC disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
 CC modulators are useful for the treatment of cancer, particularly colon
 CC cancer or cellular proliferation and/or differentiation disorders. Other
 CC disorders associated with 3714, 16742, 23546 or 13887 expression or
 CC activity that can be treated include bone related disorders, inflammatory
 CC disorders, autoimmune diseases, cardiovascular disorders and liver
 CC diseases. The present sequence represents a human protein kinase
 CC polypeptide 3714.

XX SQ Sequence 783 AA;
 Query Match 20.2%; Score 831.5; DB 22; Length 783;
 Best Local Similarity 30.3%; Pred. No. 1.2e-47;
 Matches 249; Conservative 99; Mismatches 281; Indels 193; Gaps 26;

QY 31 VGPYRLKTEKLGKQGTGLKGVHCITGOKVAIKIVNEKLSSEYIAMVEREITALKLEH 90
 DB 24 VGFYDIERTLGKGNFVAVKLARHRTVTKYAIKIDTRLDSSLEKIYEVQMLKLNH 83
 QY 91 PNYLKLHDYIENKKIYLVLEHVSGLFEDYLVKKGLTPREARKFRQIVSALDFCHSY 150
 DB 84 PHIKLYQWETMDMYIYVTEFAKNGEMFDYLSNCHLSENAKKKFWQILSAVEYCHD 143
 QY 151 SICHRDLKPEENLLDEKNNRIADFGNASIQVDSLSLETGSGPHYACPEYIKKEKDG 210
 DB 144 HLYRDLKTEENLLDGMNDIKLADFGNGNFYKSGEPLSTWCSSPVYAPAEVFEKEVEGP 203
 QY 211 RADMWSGVILFALVGALEPFDDNLRLQLEKYRGVFNHPIPPCCQSLIRGMEYEP 270
 DB 204 QLDIWSLGVVLYLVYCGSLPFDGPNLPTLRQVLEGRFRIFPMSCQCESTIRMLVYDP 263
 QY 271 EKRLSEIQKHPWYLGKHEPDCLEPAG-RRVAMRSLPSN-GEIDPVLESMAISGC 328
 DB 264 ARRTIAQIRHRRM-----RAEPCL-PGPACPAFSAHSTYSLNGDYDEQALGIMQTLY 317
 QY 329 FRDERLHRELRSEENOEKMIYLLDRKERYPSCEDDLPRLNDVDPKRKRDSMLS 388
 DB 318 --DROKTVSLSNNSYNNRAITYLLERKEYRNAD-----CA 354
 QY 389 RRG-KRRPERKMEVLSTIDAGGSPVPTRRALEMAQHSQRSVSGASTG--LSSS-- 443
 DB 355 RGPAPQPRPSDLSGLEVPQEGSLTDPFRPALCPQPTIVSQVLAEMDELQSLQ 414
 QY 444 -----PLSSPRSPVSEFSF-----EPGAGDARGGGSTKTQTLPSGPRGCGAGEQ 491
 DB 415 WPLFPFVDASCGVFRPRPVSPSSLLDTALISEARQGPGLSEEDQDES--PSTGRHRT 473
 QY 492 PPPSARSTPLGP-----PGSPRSSGGTPLH----- 518
 DB 474 LAEVSRLSPFLAPCLIVBSPSTTASPAEGSSDCLTFSAKSPAGLSGTPAIOGLIGAC 533
 QY 519 SPLHPPRASP-TGTPGTTPPSPGGVGGA-----AMSRRLN 554
 DB 534 SPVRL--ASPFILGOSQATVLAQOGLGAVLLPVSEFGRRASDPSLTGGLAFRQQLR 591
 QY 555 SIRNS--FLGSPRRH---RRKMOVPTAEEMSS-----LTPESSPEL-----AKRSMFGN 598
 DB 592 KTRTKGFLGLNKIKGLARQVCPVSPASRASRGSLSPFHAQSPGLHGAAGSREGW--- 648
 QY 599 FLSLDKEQIFVLKDKPLSIKADIVHAFSLPSLSHVSLSQTS- FRAEYKASGGSPV 657
 DB 649 -----SLEEVLEEQRLQLQHHPAAAP----- 671
 QY 658 QKPRFOVDISSEGEPPRR-----DSGGGGIVSVTF---LISGPRRRFRKRV 707
 DB 672 -----GCSQAPQAPAPAPVYIAPCDGPAAPLPSTLLTSGLPPLPPLQLGTGASPV 721
 QY 708 TIOAQLSTHDPQSVQALADEKNGAQTREPAGAPRRSLQPPG 749
 DB 722 ASAAQLDTH-----LHIGTGPAALPAVPPPRRLARLAPG 755

RESULT 12
 AAB85786
 ID AAB85786 standard; Protein; 783 AA.
 AC AAB85786;
 DT 29-OCT-2001 (first entry)
 DE Human kinase PKIN-5.

XX PKIN, kinase; cytosolic; immunosuppressive; immunostimulant; human;
 KW antiarteriosclerotic; cardiatic; gene therapy; antisense therapy.
 XX Homo sapiens.
 XX MO200160991-A2.
 XX 23-AUG-2001.
 XX 16-FEB-2001; 2001WO-US05240.
 XX 17-FEB-2000; 2000US-0183682.
 XX 02-MAR-2000; 2000US-0186559.
 XX 09-MAR-2000; 2000US-0188606.
 XX 17-MAR-2000; 2000US-0189998.
 XX 30-MAR-2000; 2000US-0193851.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Tang YF, Buford N, Gandhi AR, Patterson C, Khan FA, Yue H;
 PI Hafila A, Shih LL, Tribouley CM, Yao MG, Burrill JD, Marcus GA;
 PI Zingler KA, Lu DM, Bandman O, Pollock JL, Griffin JA, Thornton M;
 PI Nguyen DB, Lal P, Walsh RT;
 XX MPI: 2001-514771/56.
 XX N-PSDB; AAH76213.
 XX Isolated human kinase polypeptides useful in the diagnosis, treatment
 PT and prevention of cancer, immune disorders and disorders affecting
 PT growth and development -
 XX Claim 1: Page 107-109; 126pp; English.
 XX The invention provides human kinases (PKIN) and polynucleotides encoding
 CC PKIN. The PKIN polypeptides can be expressed using standard recombinant
 CC methodology. The PKIN polypeptides, polynucleotides, modulators and
 CC specific antibodies are useful in the diagnosis, treatment and prevention
 CC of cancer, immune disorders, disorders affecting growth and development,
 CC arteriosclerosis, and other cardiovascular diseases, and lipid disorders
 CC and in the assessment of the effects of exogenous compounds on the
 CC expression of nucleic acid sequences of human kinases. The present
 CC sequence represents a human PKIN-5 polypeptide.
 XX Sequence 783 AA:
 SQ
 Query Match 20.2%; Score 831.5; DB 22; Length 783;
 Best Local Similarity 30.3%; Pred. No. 1.2e-47;
 Matches 249; Conservative 99; Mismatches 281; Indels 193; Gaps 26;

OY 389 RHG-KRRPERKSMELVITDAGGGSPVTRALEMAQHORSRSVSGATG--LSSS-- 443
 DB 355 RPPANQPRRRSDLSGLXVQEBLSTDPFRALLCPQPLVSGVLAENDCELOSLQ 414
 OY 444 -----PLSFRSPVFSFSP-----EPGAGDEARGGGSPSTKOTLPSSRPGGAGEQ 491
 DB 415 WPLFFPYDAGSGVFRFRPVSPSSLDLTALISEAROGGEEEDTQESTL-PSSTGRHT 473
 OY 492 PPPPSARSTPLPGP-----PGSPRSSGCTULH----- 518
 DB 474 LAEVSTRSLPPLTAPCTIVSPSTTASPAGTSSDSCLTFPSASKSPAGLSGTATOGGLGAC 533
 OY 519 SPLHTPRASP-TGTPGTTPPPSPGGVGGA-----AMRSRLN 554
 DB 534 SPVRL--ASPLDGSQSAFTPVLAQGGIGGANLVVSPQEGRRASDTSLTQGLAKPRQDLR 591
 OY 555 SIKNS--FLGSPRPH--RRKQVPTAEEMSS-----LTPESSPEL-----AKRSWFGN 598
 DB 592 KTRIRGFLDLNLIKGLARVCQVPASRASRGSLSPHAPAQSPGLHAGAGSREGW--- 648
 OY 599 FISLDEEQIFVLKDKPLSIRADIYHAFISLPSLSHVSLSQTS-PAEKKASGGEVF 657
 DB 649 -----SLDEVELEQQRLLQLQHHPAAP--- 671
 OY 658 QKEVREQVDSISSEGEPEPSRR-----DSGGGGIYSVTFT-----LISGSRFRKRYVE 707
 DB 672 -----GCSQAPQAPAPAFVIAPCDGPGAPLSTLSGLPPLPPILLQTGASPV 721
 OY 708 TIOAQLSTHDQPSVQALADEKNGAOTRPAAPRSLQPPG 749
 DB 722 ASNAQLDTH-----LHIGTPTALPAVPPRLRLAPG 755
 RESULT 13
 AAM78349
 ID AAM78349 standard; Protein; 768 AA.
 AC AAM78349;
 XX 06-NOV-2001 (first entry)
 DE Human protein SEQ ID NO 1011.
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 OS Homo sapiens.
 XX WO200157190-A2.
 XX 09-AUG-2001.
 XX 05-FEB-2001; 2001WO-US04098.
 XX 03-FEB-2000; 2000US-0496914.
 XX 27-APR-2000; 2000US-0560875.
 XX 20-JUN-2000; 2000US-0598075.
 XX 19-JUL-2000; 2000US-0620325.
 XX 01-SEP-2000; 2000US-0654936.
 XX 15-SEP-2000; 2000US-0663561.
 XX 20-OCT-2000; 2000US-0693325.
 XX 30-NOV-2000; 2000US-0728422.
 XX (HYSE-) HYSEQ INC.
 XX Tang YF, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX MPI: 2001-476283/51.

DR N-PSDB; AAK51482.
 XX Nucleic acids encoding polypeptides with cytokine-like activities,
 PR useful in diagnosis and gene therapy -
 XX
 PS Claim 20; Page 3241-3243; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAK80303-AAK80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAK80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 CC
 XX

SO Sequence 768 AA;

Query Match 20.2%; Score 828.5; DB 22; Length 768;

Best Local Similarity 31.1%; Pred. No. 1.9e-47;

Matches 247; Conservative 110; Mismatches 260; Indels 177; Gaps 26;

QY 30 YVGPYRLKTLGGQGLVGVGCTGOKVAIKIVNREKLSSEVLMKVEREIALKLE 89
 DB 40 HIGNYRLKTLGGQGLVGVGCTGOKVAIKIVNREKLSSEVLMKVEREIALKLE 99
 QY 90 HPHVLLKLDYENKKYLYLVLEHVSGLGELFDYLVKGRLLPKEARKFPQIVSALDFCHS 149
 DB 100 HPHVLLKLDYENKKYLYLVLEHVSGLGELFDYLVKGRLLPKEARKFPQIVSALDFCHS 159
 QY 150 YSICHRPLKPEENLLDKNNIRIDFGMASIQVDSLETSCGSHYACPEYIKKEKIDG 209
 DB 160 KFTVHRPLKPEENLLDKNNIRIDFGMASIQVDSLETSCGSHYACPEYIKKEKIDG 219
 QY 210 RRADMMSGCVILFALVGLVGLPDDNDLRLEKVRGVFHHPHFPPCOSLLRMIEVE 269
 DB 220 PVDVMSGLVILYLVGSLPFDGONLKEKRLERYIKRIRPFYISTOCENIKKFLILN 279
 QY 270 PEKRLLEQIOKHMYLGKHEPD---PCLBPARGRYAMRSLPENGELDPVLESMSLI 326
 DB 280 PSKRGTEQIMKDRMMNVG--HEDELMKPYVEPLPDYK-----DPRTELMVSM 326
 QY 327 GCRDRRLRLRELSEENQEKMTIYLLDKKE----- 359
 DB 327 G---YTRREIDSLVGORYN-EVMATYLLGLKSELEGDITTLKRPASADLTNSSAHPH 383
 QY 360 -RXPSC-----EPODLP-----RNDVDPKRVDS----- 384
 DB 384 TRRSVANSKORRFSQOAGPPIPTNSYSKKTOSNNAENKRPEDRESGRASSTAKYV 443
 QY 385 ---PMLSRGKRRPERKSMSEVLSITDAGGSGSPVTRALEMA--QHSQKRSVSG--A 436
 DB 444 ASPLPGLERK-KTPTPTSTNSVLTSTNRSRNSPLERASIGQASIQNGKDSLTPMGSPA 502
 QY 437 STLSSPSPLSPR-----SVFSPSPGAGDEARGGSGSTKQTQLPSRKP----- 483
 DB 503 STASASAAVAASARROHOKMSASVHPKASGLPPTESNCEVPRPSTAPORVVASPSAH 562
 QY 484 ---RGGAGAGG---PPPSASRSTPLPDPGSPRSGSGTPLSLHTPPASPTGTGTPPP 537
 DB 563 NISSSGCAPRTPNPRGVSSKSTPHAGQLROYROQNLPT-----GVTP 606
 QY 538 PSBGGVGAAMSRSLNRSFLSGSPFRHRRKMQVPTAE--MSLTP----- 584
 DB 607 ASPSGHSQGR--RGASGSIFSCK--TSKFEVRNRLNEPESKDRVETLRHHVVGSGGNDKX 662
 QY 585 ---ESSPELAKNSW-RGNISLDEKQIPLVLDKPLSLIKADIIVHAFSLTSPSLSHSVL 639

DB 663 EEFREAKPRSLRFTWSMKTTSSMEPNEMERIRKRVLDANSQCSLHKKYML--CMGTTP 720
 QY 640 SQTSFRAEYKASGSPVFOKPVRFVDDISSEGEPEPRDRGGGGGISTFTLLISGPS 699
 DB 721 GHEDE-----VOMEMEVC-----LPRSLNSG-----VFRIRISGTS 752
 QY 700 RREKRVETIOAL 713
 DB 753 MARKNIASKIANEL 766

RESULT 14

AAE19050 standard; Protein; 691 AA.

AC AAE19050;
 DT 18-JUN-2002 (first entry)
 XX
 DE Human PAR-1B alpha protein.
 XX

KW Human; Dishevelled associated kinase; PAR-1 kinase; enzyme; PAR-1A;
 KW PAR-1B alpha; PAR-1B beta; PAR-1C; cancer; hyperproliferative disease;
 KW antisense therapy; Dsh.
 XX

OS Homo sapiens.

PN MO200210402-A2.

PD 07-FEB-2002.

PF 30-JUL-2001; 2001WO-US233981.

PR 28-JUL-2000; 2000US-221860P.

PA (CHIR) CHIRON CORP.

PI Sun T, Feng J, Reinhard C, Pantl WJ, Williams LV;

DR MPI; 2002-206192/26.

DR N-PSDB; AAD30397.

PT Isolated nucleic acids encoding the human and Drosophila Dishevelled
 PT associated kinase, referred to as PAR-1 kinase, useful for identifying
 PT modulators which can be used for modulating, inhibiting or preventing
 PT the growth of cancer cells -
 XX

PS Claim 7; Page 283-284; 287pp; English.

XX
 CC The invention relates to Dishevelled (Dsh) associated kinase, referred
 CC to as PAR-1 kinase in Drosophila and human homologues of PAR-1 referred
 CC to as PAR-1A, PAR-1B (alpha and beta) and PAR-1C. The invention also
 CC provides nucleic acid molecules encoding such proteins. PAR-1 activates
 CC the Wnt pathway and is required for Wnt signalling in mammalian cells.
 CC The PAR-1 modulators can be used to reduce the expression and/or
 CC biological activity of PAR-1. They are useful for modulating, inhibiting
 CC or preventing the growth of cancer cells. They are also useful for
 CC treating other diseases of hyperproliferation. Sequences of the invention
 CC are used in antisense therapy. The present sequence is human PAR-1B alpha
 CC protein.
 CC
 XX

SO Sequence 691 AA;

Query Match 20.2%; Score 828; DB 23; Length 691;

Best Local Similarity 30.9%; Pred. No. 1.8e-47;

Matches 232; Conservative 115; Mismatches 261; Indels 142; Gaps 24;

QY 30 YVGPYRLKTLGGQGLVGVGCTGOKVAIKIVNREKLSSEVLMKVEREIALKLE 89
 DB 16 HIGNYRLKTLGGQGLVGVGCTGOKVAIKIVNREKLSSEVLMKVEREIALKLE 99
 QY 90 HPHVLLKLDYENKKYLYLVLEHVSGLGELFDYLVKGRLLPKEARKFPQIVSALDFCHS 149

```

DB 76 HPIVILFEVIEIEKTYLWMEYASGGEVFDYLAHGMKEKEARAKROYISAVOYCHQ 135
QY 150 YSICHDLKPEKLLDKENIRIADFGMASIQVDSILETSCGSPHYACPEVYKCEKYG 209
DB 136 KFIYHDLKAEENLLDDAMNKKIADFGSNEFTFGNKLDTGCGSPHYAPAEFGCKKYG 195
QY 210 RRDMMSCGTYLALVALGALPDEDDNLRQLEKRYGFHHPHFPCCQSLRGMIYE 269
DB 196 PEVDVMSGLVILTYLVGSLPFDGQNLKELREYLRGKYRIPFYSTOCENILKFLILN 255
QY 270 PEKRLSLEJOIKHPWYVYGGKHEPD---PCLEPAPGRVAMSLPENGELDPVLSMASL 326
DB 256 PSKRGTLEQIMKRMAMVNG--HEDELKPYVPLPDYK-----DPRTELAMSM 302
QY 327 GCFRDRERLRELRSEBNOEKMTYLLDRKERYPSCEDDLPBRNDVDPREKRVDP- 385
DB 303 G--YTRIEDIDSLVGRYN--EVMATYLLGYKSELEGDITLKRPRPADLNNSSAPSS 359
QY 386 -MLSRHGKRPERKSMVLTITDAGGGSPVTRALE---MAQHSQSRSVSGASTGLIS 441
DB 360 HKVORSVSANPKORF-----SDQAGPALPTNSYSKKTQSNNAENKRPEDRESGRK 412
QY 442 SSPSSPSRVSFSPSPPGAGDEARGGGPTKTQTLPSRGRGSGAGEQPPPSARSP 501
DB 413 AS--STAKVPA---SLPGL---ERKKTPTPTNSVLTSTN-----RSRNSP 453
QY 502 L--PGPPSGRSGGTPDLSPATPRASP-----GTPGT----- 535
DB 454 LBERASLGQASIQNGKDSAPQVVPVAPSAHNISSGADRTNFPVGSRSSTFHAGQ 513
QY 536 -----PPSPRGVGGAAMRSRLNSTRNSFLSPRRHRRKMOVPTAE--MS 580
DB 514 LROVDOQNLPGVTPASPGHSGR--RGASGISFSE--TSKVRNRLNMEPSKDRYE 569
QY 581 SLTP-----ESSPELAKRW--FGNFIIDKEEQIFLYLKDPLSIRAD 623
DB 570 TLNPHYVSGGNDKEKEFPEAKRSLRFTWSMTTSMENEMREIRKYVDANSQSG 629
QY 624 IYVAFSLPBLSHSVLSQTSFRAEYKASGGSPVOKPYRFOVDISSSGPEPSPRDGS 683
DB 630 LHKRYML--CMGTGPHEDF-----VQMEMEYCK-----LPRLSLNG 665
QY 684 GGGIYVFTLLSGPSRRFRVETIAQL 713
DB 666 -----VRFKRISGTSMAFKNISKIANEL 689

RESULT 15
AAE19051
ID AAE19051 standard; protein; 724 AA.
XX
AC AAE19051;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human PAR-1B beta protein.
XX
KW Human; Dishevelled associated kinase; PAR-1 kinase; enzyme; PAR-1A;
KW PAR-1B alpha; PAR-1B beta; PAR-1C; cancer; hyperproliferative disease;
KW antisense therapy; Dsh.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH MISC-difference 274 /note- "Encoded by AGG of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 275 /note- "Encoded by TGC of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 276 /note- "Encoded by CTG of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT /note- "Encoded by CTG of the inverse complementary

```

```

FT MISC-difference 277 strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT /note- "Encoded by ACA of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 278 /note- "Encoded by CTT of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 279 /note- "Encoded by TTG of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 280 /note- "Encoded by GAC of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 281 /note- "Encoded by GAA of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 282 /note- "Encoded by GAG of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 283 /note- "Encoded by TTT of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 284 /note- "Encoded by AAA of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 285 /note- "Encoded by GAG of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 286 /note- "Encoded by TAA of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 287 /note- "Encoded by GAA of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 288 /note- "Encoded by TTA of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 289 /note- "Encoded by GGG of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 290 /note- "Encoded by TTC of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 291 /note- "Encoded by TTT of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 292 /note- "Encoded by TCT of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 293 /note- "Encoded by CCG of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 294 /note- "Encoded by TGA of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT MISC-difference 295 /note- "Encoded by AAT of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT WO200210402-A2.
XX
XX 07-FEB-2002.
XX
XX 30-JUL-2001; 2001MO-US23981.
XX
XX 28-JUL-2000; 2000US-221860P.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Sun T, Feng J, Reinhard C, Pantl WJ, Williams LT.
XX WPI; 2002-206192/26.
XX N-PSDB; AAD30398, AAD33068.
XX Isolated nucleic acids encoding the human and Drosophila Dishevelled
XX associated kinase, referred to as PAR-1 kinase, useful for identifying

```


GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2003, 12:28:38 ; Search time 48 Seconds

(without alignments)
1558.179 Million cell updates/sec

Title: US-10-003-690-2

Perfect score: 4108

Sequence: 1 MSSGAKGGGGSPAYHLPH.....PRGPPKLLATNGTLP 778

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: PIR.73:*
2: PIR1:*
3: PIR2:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1744	42.5	887	2 T20941	hypothetical prote
2	836	20.4	1518	2 S37928	probable purine nu
3	828	20.2	745	2 G01025	serine/threonine p
4	814	19.8	1192	2 T18611	probable serine/th
5	797	19.4	713	2 S27966	probable serine/th
6	795.5	19.4	512	1 JCI446	serine/threonine-s
7	793	19.3	401	2 B90120	SNF1-related prote
8	792.5	19.3	504	2 T10449	probable serine/th
9	791.5	19.3	1246	2 G89287	protein H39E23.1
10	788	19.2	774	2 I48609	probable serine/th
11	780.5	19.0	798	2 JCT500	gdk protein - chlc
12	767	18.7	1398	2 T13741	hypothetical prote
13	766.5	18.7	504	2 T07415	probable serine/th
14	766.5	18.7	511	1 A36009	serine/threonine-s
15	761.5	18.5	512	2 T52633	serine/threonine-s
16	760.5	18.5	512	2 S59359	GIN4 protein - yea
17	757.5	18.4	576	2 T41587	probable carbon ca
18	755.5	18.4	562	2 T29858	hypothetical prote
19	753.5	18.3	633	1 A26030	serine/threonine-s
20	753.5	18.3	775	2 T38929	changed division r
21	751.5	18.3	512	2 T07788	probable serine/th
22	741	18.0	473	1 S59941	serine/threonine-s
23	731.5	17.8	513	1 S60303	SMF-related kinase
24	730.5	17.6	472	2 B90100	probable protein k
25	722	17.6	915	2 S74283	probable protein k
26	719.5	17.5	602	2 S72513	FG2 protein - yea
27	718	17.5	552	1 S51025	hydroxymethylglut
28	717.5	17.5	481	2 I49072	protein kinase - m
29	715	17.4	552	1 A53621	hydroxymethylglut

30	711.5	17.3	513	1 S60304	serine/threonine-s
31	695.5	16.9	510	2 T04145	serine/threonine p
32	678	16.5	726	2 T33998	hypothetical prote
33	672	16.4	891	2 T40503	protein kinase kin
34	666.5	16.2	502	1 A41361	serine/threonine-s
35	663	16.1	891	2 A38903	protein kinase 1 -
36	656.5	16.0	651	2 S52244	p69Eg3 protein - A
37	649	15.8	1558	2 T29253	hypothetical prote
38	636.5	15.5	622	1 S44859	serine/threonine-s
39	627	15.3	480	2 A86427	probable serine/th
40	626.5	15.3	441	2 C84667	probable protein k
41	624.5	15.2	442	2 T48202	protein kinase AK2
42	624	15.2	440	2 T09903	serine/threonine-s
43	623	15.2	440	2 T14736	probable serine/th
44	621	15.1	440	2 T14735	probable serine/th
45	614	14.9	445	2 T50802	serine/threonine p

ALIGNMENTS

RESULT 1

T20941
hypothetical protein F15A2.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T20941

R:Gregory, J.

submitted to the EMBL Data Library, March 1996

A:Reference number: Z19349

A:Accession: T20941

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-887 <WILD>

A:Cross-references: EMBL:270207; PIDN:CAA94127.1; GSPDB:GN00028; CESP:F15A2.6

A:Experimental source: clone F15A2

C:Genetics:

A:Gene: CESP:F15A2.6

A:Map position: X

A:Introns: 32/1; 63/3; 92/2; 139/2; 189/3; 328/2; 448/2; 516/3; 604/2; 684/3; 735/2;

Query Match 42.5%; Score 1744; DB 2; Length 887;

Best Local Similarity 50.6%; Pred. No. 8e-55;

Matches 392; Conservative 88; Mismatches 198; Indels 96; Gaps 18;

QY	24	PPQNAQYVGRPLEKTLGKGTGLVKGVCITGOKVAIKIVNREKLSSEVLMKVEREIA	83
DB	10	PVAQAOYCGPRYKLEKTLGKGTGLVKGVCITGOKVAIKIVNREKLSSEVLMKVEREIA	69
QY	84	IKLIEHPVYLKLDVYENKRYLYLVLEHVSGETLFDYLVKGRITPREARKFFROIYSA	143
DB	70	IMKLIEHPVYLKLDVYENKRYLYLVLEHVSGETLFDYLVKGRITPREARKFFROIYSA	129
QY	144	LDPCYSYCHRDPLKPELNLLDEKNNIRADFGAASQVGDLSLETSCGSHYACPEYIK	203
DB	130	LDPCAHNHCHRDPLKPELNLLDERNNIKVADFGAASQVGDLSLETSCGSHYACPEYIR	189
QY	204	GKRYDARRADMSGVILFALLVGALEPFDDNLRQLEKVRGFGVHPHFPPPCOSILR	263
DB	190	GKRYDARRADMSGVILFALLVGALEPFDDNLRQLEKVRGFGVHPHFPPPCOSILR	249
QY	264	GMEVPEPKRLSEQIQKHFWYLGKHEPDCLEPAGRRVAMSLPSNGLDVLISM	323
DB	250	AMEVDPGRKRYSLADVFKEFW-VSGTTPKADPELELPMSQVQVTHVIGEDSIDDPDLRH	308
QY	324	ASLGCRRDRERLRELRSEENQEMTYLLDRKEVYSCQD-DLPPR---NDVPPR	379
DB	309	NCLGCKKQOKLNEILSPKHNTEKMYTFLLDKRRRPPQEDDTETVLKCAQANNPPK	368
QY	380	KRYDSEMLSRHGRKRRPDRKSMELVSLTDAGGSGPYPTRALEMAQHSORSRSVSGASTG	439
DB	369	KRYDSEMLSRHGRKRRPDRKSMELVSLTDAGGSGPYPTRALEMAQHSORSRSVSGASTG	410

440 LSSPLSPSPV-----FSFSPGAGDEARGGSPSTKT-----QTLPS 480
 411 ---SPTESRSSTRDLFGSSSSGYSARAGEDRDGRSARSTSYHYYPQVADPOTIAE 467
 481 -----KPPGGGAGE-----QPPPSASTPLPGPGSPRS 511
 468 AARHVDAQERRRSRSGRSSRKRSDKASASSSSCKNDASTSSVPKRYSPSVMS 527
 512 SGGTPLHSPDLPRASPTGPTPTPPSPGCGVGGAAMRSRLNIRNSFLGSPFHRK 571
 528 ESIVVSSSTMNNSSTNSLJAGNSQTSIGTSG--PWRSKLNNIKNSFLGTPFHRK 585
 572 QVPTAEEMSLAPE--SSPELAKRWFNGF--ISLDKEOIFLYLKDKPLSSIRADIYA 627
 586 SNGTASDESDSQMIDTDLVKKSWFGLASSMSVERDTHCVQGTKLNSIKAEILRA 645
 628 FLSTPLSHSVLSQTSFRAEYK--ASGSPSVQKPVAFQVODISSEGPSPRRDG--SGC 684
 646 FLQIHELHSHSVQNGCFREYKRGKPTVGGSVFSKIGKMDIIT-----PSPQVVIAGE 699
 685 GGIYSVFTLISGSPRRFRKRVETIOALSTFDQSPVOALADEKNGA--QTRP 736
 700 TPTVYVQFVLLAGPVARRKFLVEHLISLIL-----QNSTQGRADROQQAALMVRP 748

RESULT 2

S37928

probable purine nucleotide-binding protein YKL101w - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YKL453

C:Species: Saccharomyces cerevisiae

C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 24-Sep-1999

C:Accession: S37928; S39084

R:Cheret, G.; Fukuhara, H.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Pallier, C.; Puzo

submitted to the Protein Sequence Database, March 1994

A:Reference number: S37920

A:Accession: S37928

A:Molecule type: DNA

A:Residues: 1-1518 <CHR>

A:Cross-references: EMBL:X28101; NID:9486168; PID:CAA81941.1; PID:9486169; MIPS:YKL101w

A:Experimental source: strain S288C

R:Pallier, C.; Valens, M.; Puzos, V.; Fukuhara, H.; Cheret, G.; Sor, F.; Bolotin-Fukuhara

Yeast 9, 1149-1155, 1993

A:Title: DNA sequence analysis of a 17 kb fragment of yeast chromosome XI physically loc

protein kinases.

A:Reference number: S39084; MUID:94078677; PMID:8256524

A:Accession: S39084

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-1518 <PAL>

A:Cross-references: EMBL:X71133; NID:9431205; PID:CAA50456.1; PID:9431215

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:HS11

A:Cross-references: SGD:S0001584; MIPS:YKL101w

A:Map position: 11L

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

C:Keywords: ATP; P-loop; purine nucleotide binding; serine/threonine-specific protein ki

F:79-369/Domain: protein kinase homolog <KIN>

F:79-86/Region: nucleotide-binding motif A (P-loop)

F:87-95/Region: protein kinase ATP-binding motif

F:85/Binding site: ATP/GTP (Lys) #status predicted

Query Match 20.4%; Score 836; DB 2; Length 1518;

Best Local Similarity 28.5%; Pred. No. 1e-22;

Matches 256; Conservative 135; Mismatches 261; Indels 246; Gaps 29;

31 VGPYRLKLTGKQGTGLVKGHCITGQKVAIKVNRKL-----70

DB 78 VGPWKGTGKTLGKSSGRLAKNMETQLAIAIKVPRKKAHVHCSNNGTVPNSYSSSMVT 137

71 ----SSSVLMK-----VERIATIKLEPHPHVLKLVHYENKTKLYLVLEHVS 114

DB 138 SNVSPSIASRHSNHSQTNPDYGIERTIVIMKLISHTNVMALEFWEKSELYLVLETVD 197

115 GGELELYLKRLGRLPKARKEFRQIVSALDPCHSYSLCHRLDKPENLLDEKN--NRIA 173
 198 GGELFDLYVSKRLPERAIHYFKQIVGVSCHSNICHRLDKPENLLDKNNRRIKIA 257
 174 DFGASLOYDLSLETSGSPHYACEVIKGEYDGRADMSGCVILFALLVGLALPDD 233
 258 DGMALALEPNLKLTKSCSPHYASPEIWMGRPHGSPVMSGCVILFALLVGLHLPND 317

234 DNLRLLEKVGKGVPMHPIPPDQSLRGMEVEGERKSLDELOKHW-----284

318 DNKKLLKVGSGKTQMPENLSEARDLISKIVIDPEKITTQELIKHLIKKYDLPV 377

205 -----YLGGHEBDPCLEPARGRVAMSLSPNGELDDPVLESNASLGFCDRE 333

378 NKVLKMRKDNARAGKNSND--LHLLNNVSPSYVTLHSGKEIDESTLRSQILLHGVSR 435

334 RLHRELSEENDEKITYILLDRERY-----PSCDDPLPRNDVPPRKVDPMLS 388

436 LITAKLQPMSEKLFYLLQYKORHSISLSSSEKKSATESVNEERIEVASKTAN 495

389 RRGKRPERKSMELVSLTDAGGGSPVPTRALEMAOHSO-----RSRVSAGSTGL-- 440

496 NTGLR-----SENNDYKILHSLDT--HSEDITVQNNNAITGVNTEINA 537

441 -----SSSPLSPRS-----PVFSFSPGAGDEARGGSPSTKTQT 477

538 PVLAQKSGPSINTLQPESDKAEEAVTLPPALPIPNAS-----SSRIFRNSYTSISS 590

478 LPSKPGGAGEQPPPARST-----PLPSPGSP-----RSGGTP-----LHP 520

591 RSRSLRLNSRLSLSASYSRELVHNDENMLPOLPSPSRYSLSRRRIHASPSTKTHKS 650

521 LHTPRASPTGPTPTPPSPGCGVGGAAMRSRLNSI--RNSFLGSPRRRRKQV----- 573

651 LSRKNTAAT-----VAARTLQNSAKRSLYSLSQTSKSLNINDLVF 694

574 -----PLAEEMSLTPESPE-----LAKRWFNGFIS--LDKEOIFLYLKDK-- 615

695 DDLPLSKRASEVNVKSEPHSLSDSDFTLCQIIFGNALDRILEEED-----NKKER 749

616 -----PLSTIADIVHAFSLPSLSHVSLSQTSFAEYKASGSPS--- 655

750 DTRQRONDTKSSADTFTTSDGAVSTNKENEGEYP---TYIEKQFMMSYKPSSEMGSL 805

656 ---VFQKPVAFQVODISSEGPSPRRDGSGGGIYVFTLLSGPRRRKRVETIQ-A 711

806 SPPPEK-----ENTLSSYLEQPKPR-----AALSDITNSFKMKNQDEMRITKKIQR 856

712 QLSLTHDQPSVQALADEKNGAQTRPAGAPRSLQP-----PGRPDELSSPPR 761

857 QLKKNDRPS-----PLKPIQHDLRVNSLPLNDQKPSLSIDPRR 896

RESULT 3

G01025

serine/threonine protein kinase - human

C:Species: Homo sapiens (man)

C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 16-Dec-1998

C:Accession: G01025

R:Nayyar, E.

submitted to the EMBL Data Library, April 1996

A:Reference number: H00564

A:Accession: G01025

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-745 <NAV>

A:Cross-references: EMBL:X97630; NID:g1310674

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom

F:18-271/Domain: protein kinase homolog <KIN>

Query Match 20.2%; Score 828; DB 2; Length 745;

Best Local Similarity 31.0%; Pred. No. 1.1e-22;

Db 621 SLAPPSYKSSNTTQIAQIPPLFNNRSTATSSAQPSTGTTGRKIADPKRILPLNSTA 680
 QY 544 VGGAAWRSLNIRNSFLGSPFRHRKM-----QVPAEEMSL--TP----- 584
 Db 681 VGGHTATGAVANNNGGIPSHDHQAQOQOYMNOLSTSTMMKSLINKTKPAAGTAATSSSS 740
 QY 585 -----ESSPELAKRSMFNF----- 599
 Db 741 SSSATSTAPLAKSSGQISIAPTPEVYREDDENNENQNNVYLGVGQTPSPAVQPT 800
 QY 600 ---ISLDEKQIFLVLKDKPLSSIKADIYHAFILPSLSHVSLOTSPRAEYKAGSGPS 655
 Db 801 EDATSSSPDEKQOQAKASSTPKES-KPSMHQSPMPSPQMTAMEKLKLSGSGQTGPT 859
 QY 656 VYQKVRVRYVDLSSSGSPSPRDSGGGIGYSVFLLSGPSRRKRVETIQALLS 715
 Db 860 V-----ATGPP--PQR-----ATSSQKRSATTTISAMGA 887
 QY 716 TPDQPSVQALADEKNGAQTRPAGAPPRSLDPPGRDPELSSPPRRGPPKDKKL 769
 Db 888 SSGGAAAAAATNQSLGASPSTGASSQGYHKA---PSSSSSTNPAPHQHL 937

RESULT 5

S27966

probable serine/threonine-specific protein kinase (EC 2.7.1.-) - human

N.Alternate names: protein p78

C.Species: Homo sapiens (man)

C.Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 04-Mar-2000

C.Accession: S27966

R.Maheshwari, K.K.; Som, S.; Parasa, I.

submitted to the EMBL Data Library, January 1992

A.Description: Sequence of a cDNA encoding 78kD marker protein lost in chemically induced

A.Reference number: S27966

A.Accession: S27966

A.Molecule type: mRNA

A.Residues: 1-713 <MAN>

A.Cross-references: EMBL:M80359; NID:g189511; PIDN:AA59991.1; PID:g189512

A.Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

C.Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

F:54-307/Domain: protein kinase homology <KIN>

F:62-70/Region: protein kinase ATP-binding motif

Query Match 19.4%; Score 797; DB 2; Length 713;

Best Local Similarity 32.0%; Pred. No. 1.3e-21;

Matches 237; Conservative 112; Mismatches 254; Indels 138; Gaps 28;

QY 30 YGPPYLETGLGQGVLVKLVHCTIGQVAVITVREKISEVLMKVEEITLIKLE 89
 Db 52 HIGNYRLKTIKGNFPAKVLARHIIITGREAALKITDQTLNPSLOKLFREVRIMKTLN 111
 QY 90 HPHVLKIDHYENKKYLYLVLEHVSSEGLFDYLVKKGRLLPKARKKFFQVSLADEFCHS 149
 Db 112 HPIVYLVLEIYEQKTYLXIMEYASGKVFYVAHGRMKERKSRQIVSAVOYQHO 171
 QY 150 YSICRDRLKPELILDEKNNIRIADFGMASLQVDSLLETSCGSPHYACPVYIGKEKIDG 209
 Db 172 KRIYHDLKAEENLLADMMIKITADFEFSNEFTVGKIDPFCCGSPYAAPILFQKKYDG 231
 QY 210 RRDMMSCGYLTALLVGAIPDDDNRLQLEKVRGVFHHPHFPPCCOSGLRMIEVE 269
 Db 232 PEVDVMSLVGLVTLVSGSLPPDQNLKELREYLRKGYRIPFTWSTDCENLKLFVLIN 291
 QY 270 PEKRLSTEQIKHPWYLGKHEPD---PCLPAPGRRYAMSLPSNGEL---DPDVLESM 323
 Db 292 PIRKGTLEQIMKRWIAG--HEBELKPYEPR-----ELDISDQKIDIM 335
 QY 324 ASIGCFDRERLRHRELSSEBNOE---KMTY-----YLLDKRKRYSCEDQDLPKND 374
 Db 336 VGMG-----YSOEIQSLSKMKKDEITATYTLGK-----SSEVRPSSD 376
 QY 375 VD-----PPKRYDPSPLSHRGRRRPRKSMELVSLIDAGGGSP---VPRRALLEMAQ 425

Db 377 LNSTGQSPHHKVRQSVSSQKOR-----YSDHAGGIPSVAVYPRKNSOTSTAD 426
 QY 426 HSQR-----SRVSGASTGLSSPLSPSPRPFVSSPEPGAGEAAGGSGPTSTQTLPS 480
 Db 427 SDLKEDGISRSKSTSVAGGK---IAPASPMGNASNNKADIP-----RKSSTVBS 478
 QY 481 RQPRGAGGEOPPPPSANSTPLPGPPGSPRSSGCPPLHSPLHTPRASPTGPTTTPPPSP 540
 Db 479 SNTAGGM-----TRRTYV-----CSERTDDR--HSVIQNGKENST-IPDQRPVAS 524
 QY 541 GGGVGAAMRSL-----NSIRNSFLGSPFRHRKQVYPAEEMSLTPESSPELAKRS- 594
 Db 525 THSISAAATPDRIREPRTASTFHHQPR-BKRTATYNGPPASPSSHAPLSQTRSR 563
 QY 595 ---WFGNFI-ISLDEKQIFLVKDK-----KPLS-----SIK-----ADIVHAFSLIP 632
 Db 584 GSTIFSKITSLTIRSNVSAKQDENEKAKRSLRFTSMKRTSSMDGDMRETRKYL 643
 QY 633 SLHSVLQTSFRAEYKASGSPVYQKRVYQVYDSSSGSPSPRDSGGGIGYSVTF 692
 Db 644 DANNCDEQRE-REFLLFCVHGDAENLVQWEMEVCK-----LPLRLSLNG-----VRF 690
 QY 693 TLISGSPRRKRVETIQAL 713
 Db 691 KRISGTSLAFKNIAKIANEL 711

RESULT 6

JCI446

serine/threonine-specific protein kinase (EC 2.7.1.-) AK21 - Arabidopsis thaliana

N.Alternate names: protein Kinase SNF1 homolog

C.Species: Arabidopsis thaliana (mouse-ear cress)

C.Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 11-Jun-1999

C.Accession: JCI446; S58266; S66334

R.Lieguen, L.; Thomas, M.; Bianchi, M.; Halford, N.G.; Kreis, M.

Gene 120, 249-254, 1992

Article: Structure and expression of a gene from Arabidopsis thaliana encoding a prot

A.Reference number: JCI446; PMID:93013041; PMID:1339373

A.Accession: JCI446

A.Molecule type: DNA

A.Residues: 1-512 <LEG>

A.Cross-references: GB:M93023; NID:g166599; PIDN:AAA32736.1; PID:g166600

R.Thiemmler, F.; Kirchner, M.; Teuber, R.; Dittlich, P.

submitted to the EMBL Data Library, May 1995

A.Description: Differential accumulation of the transcripts of 22 novel protein kinas

A.Reference number: S58266

A.Accession: S58266

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 144-198 <THD>

A.Cross-references: EMBL:X86966; NID:g928909; PIDN:CAA60529.1; PID:g928910

R.Thiemmler, F.; Kirchner, M.; Teuber, R.; Dittlich, P.

Plant Mol. Biol. 29, 551-565, 1995

Article: Differential accumulation of the transcripts of 22 novel protein kinase gene

A.Reference number: S66334; PMID:96123233; PMID:8534852

A.Accession: S66334

A.Molecule type: DNA

A.Residues: 144-198 <THD>

A.Cross-references: EMBL:X86966; NID:g928909; PIDN:CAA60529.1; PID:g928910

C.Comment: This enzyme plays an important role in a signal transduction cascade regul

A.Genetics:

A.Gene: AK10; AK21

A.Introns: 64/1; 125/3; 186/3; 230/3; 322/3; 350/3; 396/3; 475/3

C.Function:

A.Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threo

C.Superfamily: AMP-activated protein kinase; protein kinase homology

C.Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kin

F:17-271/Domain: protein kinase homology <KIN>

F:25-33/Region: protein kinase ATP-binding motif

F:48, 67, 142, 144/Active site: Lys, Glu, Asp, Lys #status predicted

F:147, 151/Binding site: magnesium (Asn, Asp) #status predicted

Query Match

19.4%; Score 795.5; DB 1; Length 512;

Best Local Similarity 43.5%; Pred. No. 1,1e-21;
Matches 157; Conservative 73; Mismatches 116; Indels 15; Gaps 5;

QY 34 YRLKTLGKGTGLVKGVCITGOKVAIKIVNREKLSVY-MKVEREIALIKLIEHPH 92
Db 19 YKLGRTIGISGFGVKIAEHALTGKVAIKILNRRKIKMEKEEVRRRIKILRLFMHH 78
QY 93 VLKLDHYENKKYLYLVLEHVSQGLFDYLVKKGRLTPKARKFRQIYASLDCHSYSI 152
Db 79 IIRLYEVIETPTDIYLVMEVYNSGELFDYVEKGRLOEDENRNFQOIIISGVEYCHRMNV 138
QY 153 CHRDLPENLLDERKNIRIADFGMASLOVGSLSLETSGSPHYACPEYIKGEKDGRRRA 212
Db 139 VHRDLKPEMLLLDSKCNVAKIADFGSLNMRDGHFLTKSCGSNTYAPETISKLYAGPEV 198
QY 213 DMWSCGVIIFALLVGLALPDDDNRLRLLEKVKRGVFNHPHPIPPDQSLLRGMIEVEPER 272
Db 199 DVMSCGVIIFALLVGLALPDDDNRLRLLEKVKRGVFNHPHPIPPDQSLLRGMIEVEPER 258
QY 273 RLSTLQIQHHPYLGKHEPDCLEPAPGRVAMRSILPSNGLDPPVLESNASLGCFFRDR 332
Db 259 RVVIRPEIRHPMFOA--HLPRYLAVPDPD-----TVQQAARKIDEDIIQEVINMGF--DR 308
QY 333 ERLHRELSEENOEKMYIYLLDRKER---YPSCEDDLPVRNDVPKRRKVDSPMLS 388
Db 309 NHLIESLRNRTQNDGTVTYLLIDNRRFASGVLGAERQETMEGTPRHMPASVASPVSH 368
QY 389 R 389
Db 369 R 369

RESULT 7

B90120

SNF1-related protein kinase [imported] - Guillardia theta nucleomorph
C:Species: nucleomorph Guillardia theta
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C:Accession: B90120
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reif
Nature 410, 1091-1096, 2001

A:Title: The highly reduced genome of an enslaved algal nucleus.
A:Reference number: A99082; M01D:11323671; PMID:11323671
A:Accession: B90120
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-401 <DND>
A:Cross-references: GB:AF083031; NID:913794278; P1DN:AAK39655.1; GSPDB:GN00152
A:Gene: Kln(gnf2)
A:Map position: 3
A:Genome: nucleomorph
C:Keywords: nucleomorph

Query Match 19.3%; Score 793; DB 2; Length 401;
Best Local Similarity 41.8%; Pred. No. 1,1e-21;
Matches 157; Conservative 92; Mismatches 97; Indels 30; Gaps 8;

QY 29 QVGVYRLKTLGKGTGLVKGVCITGOKVAIKIVNREKLSV-MKVEREIALIKLIEHPH 86
Db 10 RRVGSYIILDKTIGVATGKVAIKLAKHRTKEKGVKTIKDLFDYKPSLRKIQEISVMK 69
QY 87 LIEHPHVLKLDHYENKKYLYLVLEHVSQGLFDYLVKKGRLTPKARKFRQIYASLDCHSYSI 146
Db 70 LMFHPIHVIKIVLVLEHVSQGLFDYLVKKGRLTPKARKFRQIYASLDCHSYSI 129
QY 147 CHSYSGHRLDKPENLLDERKNIRIADFGMASLOVGSLSLETSGSPHYACPEYIKGEK 206
Db 130 CHKHRTICRDLKLEMLLDMKQIKIADFGMASLSPNMLTKTFCSPHYASPEVSNP 189
QY 207 YDGRADMMSCGVIIFALLVGLALPDD--DNLRLLEKVKRGVFNHPHPIPPDQSLLRG 264
Db 190 YNGIKADIMSGIILYLVVGLALPDDENDNRRKLEKIRFEPRPRIRIHPDRCRLIRA 249

QY 265 MIEVEPERKLSLQIQHHPYLGKHEPDCLEPAPGRVAMRSIP-----SNGELD 316
Db 250 LRLTNPKRITIDIKKHPYKSSP-LDESC-----RISTONISFASVQNPISNP--D 299

QY 317 PDVLESNASLGCFFDRRLHRELSEENOEKMYIYLLDRKERYPSCEDDLPVRNDV 376
Db 300 PETISFLPIQLVLDKDKTLGILSTKPSFLRVLYRQLEMRKMKNDVSRNLI-----NEAN 356
QY 377 PPRRRVSPMLSRGK 392
Db 357 FKRRK-----LFQKGK 367

RESULT 8

T10449

Probable serine/threonine-specific protein kinase (EC 2.7.1.-) - cucumber
N:Alternate names: SNF1-related protein kinase
C:Species: Cucumis sativus (cucumber)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
C:Accession: T10449
R:Gumpel, N.J.
submitted to the EMBL Data Library, December 1996

A:Reference number: Z17020
A:Accession: T10449
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-504 <GUM>
A:Cross-references: EMBL:Y10036
A:Experimental source: cv. Masterpiece; cotyledon
C:Function:
A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threo
C:Superfamily: AMP-activated protein kinase; protein kinase homology
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:6-260/Domain: protein kinase homology <KIN>

Query Match 19.3%; Score 792.5; DB 2; Length 504;
Best Local Similarity 45.6%; Pred. No. 1,4e-21;
Matches 151; Conservative 68; Mismatches 101; Indels 11; Gaps 4;

QY 34 YRLKTLGKGTGLVKGVCITGOKVAIKIVNREKLSV-MKVEREIALIKLIEHPH 92
Db 8 YKLGRTIGISGFGVKIAEHALTGKVAIKILNRRKIKMEKEEVRRRIKILRLFMHH 67
QY 93 VLKLDHYENKKYLYLVLEHVSQGLFDYLVKKGRLTPKARKFRQIYASLDCHSYSI 152
Db 68 IIRLYEVIETPTDIYLVMEVYNSGELFDYVEKGRLOEDENRNFQOIIISGVEYCHRMNV 127
QY 153 CHRDLPENLLDERKNIRIADFGMASLOVGSLSLETSGSPHYACPEYIKGEKDGRRRA 212
Db 128 VHRDLKPEMLLLDSKCNVAKIADFGSLNMRDGHFLTKSCGSNTYAPETISKLYAGPEV 187
QY 213 DMWSCGVIIFALLVGLALPDDDNRLRLLEKVKRGVFNHPHPIPPDQSLLRGMIEVEPER 272
Db 188 DVMSCGVIIFALLVGLALPDDDNRLRLLEKVKRGVFNHPHPIPPDQSLLRGMIEVEPER 247
QY 273 RLSTLQIQHHPYLGKHEPDCLEPAPGRVAMRSILPSNGLDPPVLESNASLGCFFRDR 332
Db 248 RITTEIRHPMFOA--HLPRYLAVPDPD-----TVQQAARKIDEDIIQEVINMGF--DR 297
QY 333 ERLHRELSEENOEKMYIYLLDRKERPS 363
Db 298 NQVLESIRNRIONKATVATYLLIDNRRFVS 328

RESULT 9

G89287

protein H39E23.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: G89287
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998

Db 500 SLTPGSRSTASASAAVSAARPHQOKSMASVHPNKAISGLPTESNCEVPRPSTAPOR 559
 QY 512 SGGTPLHPLTPRASPTGCTT-----ppspg 541
 Db 560 --VPVPSAHNITSSSGADPTNFPVGSSTFHAQOLQVQDQNLPGVTPASPS 616
 QY 542 GGVGAARSLNLSIRNFIQSPFRHRRMGOVPAE--MSILTP-----584
 Db 617 GHSGR--KQASGSIFFSF--TSKFRRLNLPESKDRVETLRPHVYVSGSGTDKDEFR 672
 QY 585 ESSSELAKRSW--FGNFISLDKEEQIFLYLKDPLSSIKADIYHAFSLIPSLSHVLSOTS 643
 Db 673 EAKPRSLRFTSMKTSTSMENEMREIRKVLADANSQSELHERVLL--CVHGTPGHEH 730
 QY 644 FRAEYKASGGSVFOKPYRFOVDISSSEGPSPRRDGGGCGGIVSYTTLISGSPRRK 703
 Db 731 F-----VQMEVEVCK-----LPRLSLNG-----VRFKRISGTSMAFK 762
 QY 704 RVEFIQAOL 713
 Db 763 NIASKIANEL 772

RESULT 11

qik protein - chicken
 N/Alternate names: Qln-induced kinase
 C/Species: Gallus gallus (chicken)
 C/Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
 C/Accession: J07500
 R/Xia, Y.; Zhang, Z.; Kruse, U.; Vogt, P.K.; Li, J.
 Biochem. Biophys. Res. Commun. 276:564-570, 2000
 A/Title: The new serine-threonine kinase, Qik, is a target of the qin oncogene.
 A/Reference number: J07500
 A/Contents: Embryo fibroblasts
 A/Accession: J07500
 A/Molecule type: mRNA
 A/Residues: 1-798 <XIA>
 A/Cross-references: GB:AF219232
 C/Comment: This protein, a member of the AMPK/SNF1 family of serine/threonine kinases,
 C/Genetics:
 A/Gene: qik
 C/Keywords: protein kinase

Query Match 19.0%; Score 780.5; DB 2; Length 798;
 Best Local Similarity 29.6%; Pred. No. 5.3e-21;
 Matches 244; Conservative 105; Mismatches 265; Indels 209; Gaps 25;

QY 17 LPHPHPPPOAOVYGPYRLEKTKGQIGVLCGVCITGQKVAIKIYNREKLSVLM 76
 Db 10 VPAPSAQPRPLR--VGYDIERLTGKGNFAYVKLARHRYTKQVAIKITDKTLDPSNLE 68
 QY 77 KVERIALILKLEHPIVHLVDYENKRYLYLVLEHVSQGELEDYLVKGRILPKEARKE 136
 Db 69 KIYREVQIMKILNPHIILKLYQVEMTKDMLIVTEFAKNKGEHDLTSGHLSSESARKK 128
 QY 137 FQIYVSADFECHSISICRDLKPEMLLIDDEKNNIRIADGMSLQVDSILETSCSPHY 196
 Db 129 FQIILSAEYCHSHIIVARDIKTEMLLDANNIKLADPGFNFYKSGPILSTWCSPPY 188
 QY 197 ACPEYIKKEKDYGRADMGSGVILFALLVGLPDDDLRLQLEKVGKGVFMPIPIPP 256
 Db 189 AAPEYBEKEKEGPHLDWISGLVLYVCGSLPFDGPNLPTLRQVNLBGRFRIPIYFMS 248
 QY 257 DCQSLKQMIIEVEPEKRLSDIOIKHPYILGKABPDPCLEPAPGRVAMRSLPSN--GEL 315
 Db 249 DCEFLIRMLVYDPTKRITISQIKQHKW-----QADPSLRQOQSLSFQMNQNSMLGDY 303
 QY 316 DPVLESASIGCFPRDRRLRLRELSEENDEKMTIYLLIDKKEKPPSCEDDLPFRNDY 375
 Db 304 NEQVGLIMQTLGI--DRQRTVESLQNSSYNHFAAIYLLLEKLEKYS-----349
 QY 376 DPKRKVDYPLSLRIGKRRPER--KSMETLSTIDAGGGSPVPTLRALEMAQHSQRSRVS 424

Db 350 -----SOLSSRPATGRQQRPSSELSN-----AEKPQSLSETL- 384
 QY 435 GASTGLSSP-----ISSPRSPVFSPEPGADBARGGG--SPTSKTQTL 478
 Db 385 -RSSLLYQQPQSLQPSLQAEEMDCMNNPLQPVF--FPVDPNNGFLRNSSISPSLLETT 442
 QY 479 PSRPPRGGAGEQPPPPARSTPLPGPGSPR--SSGTPPLHPLHTP-----RASPT 529
 Db 443 ISEEVN--QEKLEDEIKAYDHPRIPIPSNTRRHRLAEVTHFYQHPAPCIYISSASPT 500
 QY 530 -GTPGTTPPS-----PGGVGAAMRSRLNSINSFLGSPRFRKKQVPT 575
 Db 501 EGTSSDSCITSSSDSSVALSSCLAQVWTSAPATAMTS--AFLAS--QSDAPVLYQVG 556
 QY 576 AEEMSLTPESSPE-----LAKRSWFGNFISLDKEEQIFLYLKDPL 616
 Db 557 CMGASILPVEFQEGRRASDSTLTGLKAFRQOLKRNARAKFLGLNK-----604
 QY 617 LSLIKADIVHAFSLIPSLSHVLSQTSFRAEYKASGGSVFOKPYRFOVDISSSEGP 676
 Db 605 -----IKGFARQVCOSSSRRAARSAM-----S 626
 QY 677 PRDGGGGGITYSVTFLISGPRRFRKRYERTI--QAQLSTHDQPSYQALADEKNGA 733
 Db 627 PFOHAQPTCTIYS-----SSGSSREGNMLEVLOOQRMLOLQHHD-----667
 QY 734 TRPAGAPRSLQPPGRPDPELSSSPRRGPPKDKLILATNGTP 776
 Db 668 -----LLQF-----ACPQTSQTSATNGIP 686

RESULT 12

hypothetical protein 22E5.8 - fruit fly (Drosophila melanogaster)
 C/Species: Drosophila melanogaster
 C/Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
 R/Murphy, L.; Harris, D.; Barrell, B.
 Submitted to the EMBL Data Library, April 1999
 A/Description: Sequencing the distal X chromosome of Drosophila melanogaster.
 A/Reference number: Z17668
 A/Accession: T113741
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1398 <MUR>
 A/Cross-references: EMBL:AL031765; NID:e1371523; PID:e1329905; PIDN:CAA21125.1
 C/Genetics:
 A/Cross-references: Flybase:FBgn0000667
 A/Intons: 205/3; 227/1; 322/3; 688/3; 782/3; 814/2; 1363/3
 A/Note: EG:22E5.8

Query Match 18.7%; Score 767; DB 2; Length 1398;
 Best Local Similarity 28.2%; Pred. No. 2.5e-20;
 Matches 251; Conservative 112; Mismatches 324; Indels 202; Gaps 24;

QY 9 GGGSPAHPLPHPHPOHA-----QYGVYRLEKTLGKQGTGLKLG 51
 Db 105 GGGST-----PGSPSTSAVAGAGISGKDLTKLEKPMVGYDIERITGKGNFAYVKIA 158
 QY 52 VHCITGQKVAIKIYNREKLSVLMKYVEREIALILKLEHPIVHLKLDYENKRYLYLVE 111
 Db 159 RHRTKNEVAIKIIDKSDQDNTLNQKYREVEIMRLKPHILIKLYQVEMTKNMIYISE 218
 QY 112 HVSQGELEDYLVKGRILPKPKARKEPFOIVSALDCHSYICHRDLKPEMLLIDDEKNNIR 171
 Db 219 YASGEIFDIYIAKGRMSESAARFEKQIISAVEICHKKGIYHRDLKPEMLLIDLMANK 278
 QY 172 IADFGMASLYQVDSILETSCSPHYACEYIKGEYKDYGRADMGSGVILFALLVGLALPF 231
 Db 279 IADGFSNHFRPGEILLATWCGSPPYAAAEVEVEGKQYTGPEIDWISGLVLYVGLALPF 338
 QY 232 DDDNLRLQLEKVKKGVFMPIPIPPDCQSLKGMILEVEPEKRLSDIOIKHPWYLGKHE 291

```

Db 339 DGSILQSLRDLVLSGRFRIPFMSSECHLIRMLVLEPTRRYTIDQIKRHHMM----- 392
QY 292 PDCLEAPGRVAMRSL-----PENGELDPVLESMA--JGFRDRERLHRLRSEENO 346
Db 393 ---CPLELHVLAKYNLGAERQTSVEPSEDILRTIMAEVYGSGDKTR--ASLKNVYDH 447
QY 347 EKMIYVLLDR---KER-----YPSCEDDLPFRNDVPFRKR-----V 382
Db 448 VAAIYLLDQRVSHKKQSNGLGASALASTSASMTIYSNRDHPQTQOQSOQSTIST 507
QY 383 DSPMLSHKRRPRKRSMEVLSITDAGGGSPV-----PTRLAEMAHSOR 429
Db 508 SSTLAKDOCKHRLSRHQTVMSERNAHAGAPTVDPDPGVYAKYGLPLPLTGHSHL 567
QY 430 SRSVSGASTGSSPLSPSPVSESPGAGDEARGGSPSK---TQTLPSNGPR-- 484
Db 568 TGYLNGGCVEDASGIPLP---MRYTLPYPAAPSPNCSTSRVGRHSLSSSPKSH 623
QY 485 -----GGGAGEOPPPPSARSTPLPG----- 504
Db 624 RPAVLSLIDNPNPLANLRCREMMAGG-----PVGAVGYPLASKQHLQTISETLIKQ 677
QY 505 -----PGSPNSSGGTPLHSPHTPPASPTGPTTPPSPGG 543
Db 678 STEDCRALLQOSTAVAEKGKDPKPAESSVGVV-----PPASTPTTSTAGPES--- 726
QY 544 VGGAAMRSRLNLSINFLSGFRFRKMQVPTAEEMSLTPRESSPELAKRSFNGFNISLD 603
Db 727 -GSPPCGEING-----KTKMKSSSSSFDSKAMNQST--RKMSA 765
QY 604 KEQIPLVLKDKPLSIKADIVHAFSLPSLSHVLSTQSFRAEKASGSPVFOKPVRF 663
Db 766 EASLFTGLDESLPVEGRTRKRVHVGSTNGSGDGGQETNDK---SNGDSRSEKVLVA 822
QY 664 QVDISSEGEPPSPRRD-----GSGGGGIVSVTTLISGPRRRKRVETIQAOQL 714
Db 823 QGSSSTDEGCTDGCNPNPGSASQESKSGGSGNANGGPFSSHSSDRLRLGLVTASG-- 880
QY 715 SHDQPSVQALADEKNGQTRPAGAPRSLOPPGCRPDPELSSSPRRGP 763
Db 881 QSHKMR-SYASSSSSGVLAGSAGSYKSLSONLSRGS---SKMCSGP 925

```

RESULT 13

T07415
 probable serine/threonine-specific protein kinase (EC 2.7.1.1) PKIN1 - potato
 N:Alternate names: SNF1-related protein kinase
 C:Species: Solanum tuberosum (potato)
 C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 20-Jun-2000
 C:Accession: T07415
 R:Halford, N.G.
 submitted to the EMBL Data Library, March 1996
 A:Reference number: Z16022
 A:Accession: T07415
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Residues: 1-504 <HAL>
 A:Molecule type: DNA
 A:Cross-references: EMBL:X95997; NID:91216279; PIDN:CAA65244.1
 C:Gene: PKIN1
 A:Gene: PKIN1
 A:introns: 62/1; 123/3; 184/3; 228/3; 290/3; 320/3; 348/3; 389/3; 466/3
 C:Function:
 A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
 C:Superfamily: AMP-activated protein kinase; protein kinase homology
 C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
 F:15-269/Domain: protein kinase homology <KIN>

Query Match 18.7%; Score 766.5; DB 2; Length 504;
 Best Local Similarity 41.1%; Pred. No. 1.1e-20;
 Matches 157; Conservative 74; Mismatches 124; Indels 27; Gaps 6;
 QY 30 YGPFRLKTLGKQGTGLVKGVCITGQKVAIKVIREKT-SESVLKVEREIALKLII 88

```

Db 13 YLKNYRVKGLTGHGSEFGKVAIEHLTGHKVAIKILNRKKKTPDMEEKLRREIKICLTF 72
QY 89 EHHVYVLTQHYENKKTLYLVLEVSGSELDYLVKGRGLTPKARKFRQIVSLDCH 148
Db 73 VHPHVRLVLEVEITPDYIVMEYVKSGELPDYIVEGRLQEDARKTFQDIIAGVECH 132
QY 149 SYSICHRDLKPENLIDDEKNNIRIADFGMASLQVDSILLETSCGSPHYACPEVJKEGYD 208
Db 133 KNNVYHDLKPENLIDARNRKLTADFGIMDMGHFLTKSCGSPHYAAPEVYSKITYA 192
QY 209 GRADWMSGVILFALLVGLPFDNDNLKQLLEKVKGVFHPHTIPDCSLRGMTLV 268
Db 193 GPVEDVWMSGVIIYALLCGLTPEFDENIPLNFKKISGVYTLPSHLSPDLRLPRLIV 252
QY 269 EPEKRLSELDIQHPWVLGKQKHEPDLERAPGRVAMRSLPENGELDPVLESMAISGC 328
Db 253 DPKKRISVPIRQHQWF--KILPRYLAVPPDARQHLK-----KIDEELIQVSRMGL 304
QY 329 FDRERLHRLRSEENOEMKIYVLLDRKRRYS-----CEDDLPFRNDVD 376
Db 305 --DRDQLDLSLRIDQDAVAYVLLVDRNMASSGYLGAEGESYDCYSGILFPNDL- 361
QY 377 PPKRVDSPLSRHGRKRRPERK 398
Db 362 ---QLSTGNGVSESLRRPFRK 380

```

RESULT 14

A56009
 serine/threonine-specific protein kinase (EC 2.7.1.1) NPK5 - common tobacco
 C:Species: Nicotiana tabacum (common tobacco)
 C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 16-Jun-2000
 C:Accession: A56009
 R:Murataka, T.; Banno, H.; Machida, Y.
 Mol. Cell. Biol. 14, 2958-2965, 1994
 A:Title: Characterization of tobacco protein kinase NPK5, a homolog of Saccharomyces
 tase of Saccharomyces cerevisiae.
 A:Reference number: A56009; MID:94217693; PMID:8164654
 A:Accession: A56009
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-511 <MUR>
 A:Cross-references: GB:D26602; NID:9496384; PIDN:BA05649.1; PID:9496385
 C:Function:
 A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threo
 C:Superfamily: AMP-activated protein kinase; protein kinase homology
 C:Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein
 F:17-33/Region: protein kinase homology <KIN>
 F:25-33/Region: protein kinase ATP-binding motif
 F:48-67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted
 F:147,151/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 18.7%; Score 766.5; DB 1; Length 511;
 Best Local Similarity 44.4%; Pred. No. 1.1e-20;
 Matches 151; Conservative 67; Mismatches 107; Indels 15; Gaps 5;
 QY 34 YRLKTLGKQGTGLVKGVCITGQKVAIKVINEKLSSEVL-MKVEREIALKILIEPH 92
 Db 19 YKLGKTLGSGFQKVAIEHTLTGKVAIKILNRKKIKNMMEKVRREIKILRIFMPPH 78
 QY 93 VLKLDVYENKKTLYLVLEVSGSELDYLVKGRGLTPKARKFRQIVSLDCHSYSI 152
 Db 79 IILLYEVEIPSDIYVMEYVKSGELDYVKEKRLQEDARKFFQOILISVEYCHRMV 138
 QY 153 CHRDLPENLIDDEKNNIRIADFGMASLQVDSILLETSCGSPHYACPEVJKEGYDGRA 212
 Db 139 VHRDLKPENLIDSKVMVKTADFGLSNIMRDGHLTKSCGSPHYAAPEVIGKLYAGPEV 198
 QY 213 DMWMSGVILFALLVGLPFDNDNLKQLLEKVKGVFHPHTIPDCSLRGMTIEVEPEK 272
 Db 199 DVMSCGVILYALLCGLTPEFDENIPLNFKKISGVYTLPSHLSGARDLIPRLIVDPK 258

[illegible]

RESULT 15

T52633

serine/threonine-specific protein kinase (EC 2.7.1.-) AKIN11 [validated] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 24-Oct-2000 #sequence_revision 24-Oct-2000
C;Accession: T52633

R; Bhalariao, R. P.; Salchert, K.; Bako, L.; Okresz, L.; Szabados, L.; Muranak

Proc. Natl. Acad. Sci. U.S.A. 96, 5322-7, 1999

A; Reference number: Z25116; MUID:99238528; PMID:10220464

A:Status: Preliminary: translated from CB/EWBR

A;Molecule type: mRNA

A;Residues: 1-512 <BHA>
A;Cross-references: EMRL:X99279. PIDN:CAA67671 1

A; Experimental source: cultivar Columbia

A:Gene: AKT11
c;genetics:

C;Function:

A;Description: EC 2.7.1 complements SNE1 mutati

C; Superfamily: AMP-activated protein kinase; pro

C; keywords: ATP; phosphotransferase; ser

Query Match

Best Local S
Matches 148

CONCEPT	NO.	MINIMUM COUNTRIES	INDICES	GAPPS	5;
1. CONCEPT	10	30	19	6	5
2. CONCEPT	10	30	19	6	5
3. CONCEPT	10	30	19	6	5
4. CONCEPT	10	30	19	6	5
5. CONCEPT	10	30	19	6	5
6. CONCEPT	10	30	19	6	5
7. CONCEPT	10	30	19	6	5
8. CONCEPT	10	30	19	6	5
9. CONCEPT	10	30	19	6	5
10. CONCEPT	10	30	19	6	5

QY 34 YRLEKTLGGGTGLVKLGVCITGQ

Db 20 YKLGKTLGIGSFCKVIAEHVVTGHKVAIKILNRRAIKNMEMEKKVRREIKILRLFMHH

93 VIKLHDVYENKKRY.I.YI.YI.FHNSGCEI.FDVI.YKKB.I.EBFEABVEEDOTVCSAT.DDOWURST.1.CC

[illegible]

DB 80 LRQYEVLETTISDIYVMEYKSGELFDYVEKGRLOEDARNFFQOIIISGVEYCH

QY 153 CHRDLPENLLDEKNNIRIADFGMASLQVGDLSLETSCGSPHYACPEVIKGEKYDGRA 212

db 140 YHBDLKPENLILDSRCNKTADGTSNVMPCHEIKTSCSPNVAADVEYTCGVVAGDMM 100

010 ***** CERNWINDOIT EN1SCCBFN1KAFEV1BGNLIAGPEV 19

213 DMWSCGVLFFALLVGALPFDDDDNLRLQLEKVKRGVFHMPHFIPDCQSLRGMIEVEPEK 272

Db 200 DWSCGVILALCGTLPEDDENIPNLFKKIKGGIYTLPSHLSSEARDLIPRLIVDPVK 259

273 RLSLEOIOKHPVYLGKHEPDCI,EPAPGRVAMBSI,PSNCF---LDDVREVA,CTCC 333

[illegible]

DB 260 RITPEIRQHRWF--QTHLP-----RYLAVSPPDTEVEQAKKINEEIVQEVVNMGE 307

QY 329 FRDRERLHRELSEENQEKMIYYLLDRKERYPs 363

Db 308 --DRNOVLESTRNNTONDAVTYYIITDNEPIYDS 340 .

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040

Search completed: April 15, 2003, 12:33:02

Job time : 53 secs

Search completed: April 16, 2003, 12:33:02
Job time : 53 secs

Job time : 53 secs

